DOCUMENT RESUME

ED 133 236 SE 021 942

AUTHOR Osborne, Alan R., Ed.

TITLE Investigations in Mathematics Education, Vol. 9 No. 3.
INSTITUTION Ohio State Univ., Columbus. Center for Science and Mathematics Education.
PUB DATE 76
NOTE 70p.; Contains occasional light type
AVAILABLE FROM Information Reference Center (ERIC/IRC), The Ohio State University, 1200 Chambers Rd., 3rd Floor, Columbus, Ohio 43212 (Subscription, $6.00, $1.75 single copy)

EDRS PRICE MF-$0.83 HC-$3.50 Plus Postage.
DESCRIPTORS *Abstracts; Curriculum; Elementary Secondary Education; Higher Education; *Instruction; Learning; *Mathematics Education; Number Concepts; Problem Solving; Research; *Research Reviews (Publications); *Teacher Education

ABSTRACT

Eighteen research reports related to mathematics education are abstracted and critically analyzed. Seven of the reports deal with general instructional concerns (feedback on homework, error patterns, teacher-student interaction, reading strategies, teacher effectiveness, school attendance, and learning environment), four with topics in mathematics instruction (use of calculators, word problems, conservation of numbers, two with predictors of success in mathematics, and two with elementary teacher education. The remaining three reports concern curriculum, research methods, and concept attainment. Research related to mathematics education which was reported in RESOURCES IN EDUCATION (RIE) and CURRENT INDEX TO JOURNALS IN EDUCATION (CIJE) between April and June 1976 is listed. (DT)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

Vol. 9, No. 3 Summer 1976

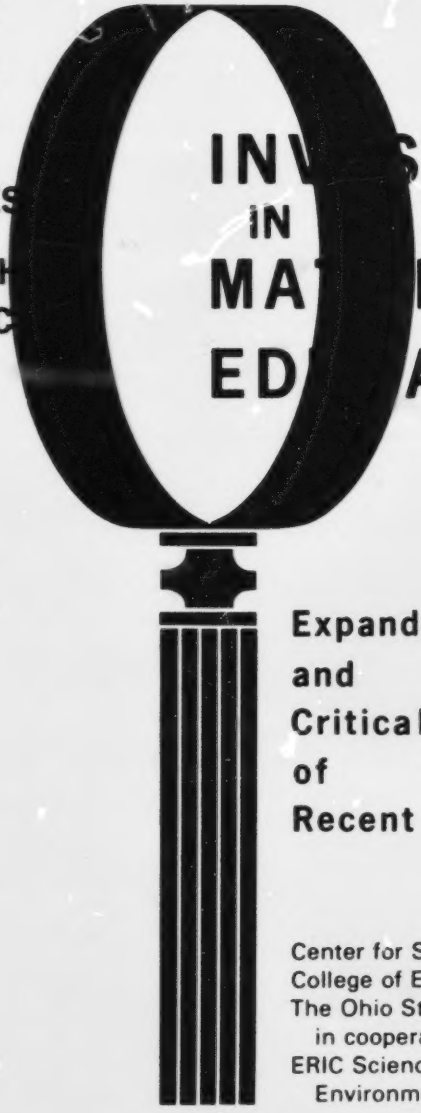
ED133236

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

INVES
IN
MATH
EDUC

INVESTIGATIONS IN MATHEMATICS EDUCATION



Expanded Abstracts and Critical Analyses of Recent Research

Center for Science and Mathematics Education
College of Education
The Ohio State University
in cooperation with
ERIC Science, Mathematics, and
Environmental Education Clearinghouse

SE 021 942

INVESTIGATIONS IN MATHEMATICS EDUCATION

Editor

Alan R. Osborne
The Ohio State University

Associate Editor

Marilyn N. Suydam
The Ohio State University

Advisory Board

E. G. Begle
Stanford University

Joseph N. Payne
University of Michigan

Len Pikaart
Ohio University

J. F. Weaver
University of Wisconsin

Published quarterly by

The Center for Science and Mathematics Education
The Ohio State University
1945 North High Street
Columbus, Ohio 43210

With the cooperation of the ERIC Science, Mathematics, and Environmental Education Information Analysis Center

Volume 9, Number 3 - Summer 1976

Subscription Price: $6.00 per year. Single Copy Price: $1.75
Add 25¢ for Canadian mailings and 50¢ for foreign mailings.

Austin, Joe D. Do Comments on Mathematics Homework Affect Student Achievement. School Science and Mathematics, 476 n2, pp159-164, February 1976.
Abstracted by BOYD D. HOLTAN 1

Buchman, Aaron L.; Hawthorne, Frank S. The Content of High School Computer Mathematics Courses. School Science and Mathematics, v75 n6, pp543-549; October 1975.
Abstracted by JANE DONNELLY GAWRONSKI 4

Cox, L. S. Systematic Errors in the Four Vertical Algorithms in Normal and Handicapped Populations. Journal for Research in Mathematics Education, v6 n4, pp202-220, November 1975.
Abstracted by RICHARD A. LESH 7

Erlwanger, Stanley. The Observation-Interview Method and Some Case Studies. Proceedings of the Conference on the Future of Mathematical Education, pp125-142. Florida State University, Project for the Mathematical Development of Children. 1975.
Abstracted by VINCENT J. GLENNON 10

Fahrmeier, Edward D. The Effect of School Attendance on Intellectual Development in Northern Nigeria. Child Development, v46 n1, pp281-285, March 1975.
Abstracted by LEWIS R. AIKEN 14

Firestone, Glenn; Brody, Nathan. Longitudinal Investigation of Teacher-Student Interactions and Their Relationship to Academic Performance. Journal of Educational Psychology, v67 n4, pp544-550, August 1975.
Abstracted by JOHN W. GREGORY 17

Forness, Steven H.; Guthrie, Donald; Hall, Robert J. Follow-up of High-Risk Children Identified in Kindergarten through Direct Classroom Observation. Psychology in the Schools, v13 n1, pp45-49, January 1976.
Abstracted by NORMAN L. WEBB 21

Fuson, Karen. The Effects on Preservice Elementary Teachers of Learning Mathematics and Means of Teaching Mathematics through the Active Manipulation of Materials. Journal for Research in Mathematics Education, v6 n1, pp51-63, January 1975.
Abstracted by KENNETH A. RETZER 25

Gaslin, William L. A Comparison of Achievement and Attitude of Students Using Conventional or Calculator-Based Algorithms for Operations on Positive Rational Numbers in Ninth-Grade General Mathematics. Journal for Research in Mathematics Education, v6 n2, pp96-108, March 1975.
Abstracted by JANE O. SWAFFORD 30

House, Peggy A. Learning Environments, Academic Self-Concepts, and Achievement in Mathematics. Journal for Research in Mathematics Education, v6 n4, pp244-252, November 1975.
Abstracted by FRANK K. LESTER, JR. 33

Klausmeier, Herbert J.; Feldman, Kathernine Voerwerk. Effects of a Definition and a Varying Number of Examples and Nonexamples on Concept Attainment. Journal of Educational Psychology, v67 n2, pp174-178, April 1975.
Abstracted by THOMAS J. COONEY 36

Linville, William J. Syntax, Vocabulary, and the Verbal Arithmetic Problem. School Science and Mathematics, v76 n2, pp152-157, February 1976
Abstracted by JEREMY KILPATRICK 39

Mayer, Richard E. Forward Transfer of Different Reading Strategies Evoked by Testlike Events in Mathematics Text. Journal of Educational Psychology, v67 n2, pp165-169, April 1975.
Abstracted by CLYDE A. WILES 42

Mpiangu, Benayame Dinzau; Gentile, J. Ronald. Is Conservation of Number a Necessary Condition for Mathematical Understanding? Journal for Research in Mathematics Education, v6 n3, pp179-192, May 1975.
Abstracted by JAMES J. HIRSTEIN 46

Richardson, Lloyd I. The Role of Strategies for Teaching Pupils to Solve Verbal Problems. Arithmetic Teacher, v22 n5, pp414-421, May 1975.
Abstracted by AARON D. BUCHANAN 49

Robitaille, David F. Criteria for Assessing the Effectiveness of Teachers of Secondary School Mathematics. Journal for Research in Mathematics Education, v6 n2, pp77-87, March 1975.
Abstracted by ARTHUR F. COXFORD 53

Telgedy, Gabriel A. The Effectiveness of Four Readiness Tests as Predictors of First Grade Achievement. Psychology in the Schools, v12 n1, pp4-11, January 1975.
Abstracted by MERLYN J. BEHR 56

Warkentin, Gary. The Effect of Mathematics Instruction Using Manipulative Models on Attitude and Achievement of Prospective Teachers. Journal for Research in Mathematics Education, v6 n2, pp88-94, March 1975.
Abstracted by LARRY SOWDER 59

Mathematics Education Research Studies Reported in Research in Education (April - June 1976) 63

Mathematics Education Research Studies Reported in Journals as Indexed by Current Index to Journals in Education (April - June 1976) 67

DO COMMENTS ON MATHEMATICS HOMEWORK AFFECT STUDENT ACHIEVEMENT. Austin, Joe D. School Science and Mathematics, v76 n2, pp159-164, February 1976.

Expanded Abstract and Analysis Prepared Especially for I.M.F. by Boyd D. Holtan, West Virginia University.*

1. Purpose

The purpose of the study was to determine the value of teacher comments on homework papers in relation to mathematics achievement.

2. Rationale

A limited number of studies have been conducted supporting the belief that regularly assigned homework is important in improving student achievement in mathematics. However, few studies have been concerned with how homework is used in the instructional process; for example, consideration of comments on homework papers and their relationship to achievement.

Previous research by Page (1958) revealed that comments on test items significantly increased student achievement. However, more recent research (Schoen and Kreye, 1974) attempted to generalize Page's study to teacher training groups and found no significant differences on test achievement.

3. Research Design and Procedure

One elementary and five secondary inservice teachers were involved in the study which consisted of nine classes with 222 students. Students were randomly divided into two groups, comment (C) and no-comment (NC). A non-parametric, two-way analysis of variance on average grades from the previous six weeks was used to test equivalency of each group of students; no significant difference was found.

For a six-week period, homework was collected by the teachers, divided by treatment group (comment or no-comment), and graded accordingly. A non-parametric analysis of variance with one degree of freedom was used to test scores on an examination administered at the end of the treatment period. The null hypothesis tested for each group was:

> There is no difference between the comment group and the no-comment group as measured by examination averages.

Statistical data for pre-experiment and examination scores reported by the researcher for each class are shown in Table 1.

---

*Assistance by Renny Azzi, WVU Graduate Assistant, Mathematics Education.

Table 1

Statistical Data for Pre-Experiment and Examination Scores

Class	Subject	Grade	Pre-experiment average R_c	R_{nc}	X_1^2	Examination scores R_c	R_{nc}	X_1^2	
1	Arithmetic	4	14.0	13.0	0.11	15.0	12.0	1.00	
2	Gen. Math	9	21.5	20.5	0.71	26.0	16.0	7.14	**
3	Gen. Math	9-10	15.0	15.0	0.00	16.0	14.0	0.40	
4	Gen. Math	9-10	14.0	13.0	0.11	13.0	14.0	0.11	
5	Algebra I	9	22.0	23.0	0.07	23.0	22.0	0.07	
6	Algebra I	9	15.5	20.5	2.08	18.0	18.0	0.00	
7	Algebra I	9	18.0	21.0	0.69	19.0	20.0	0.08	
8	Algebra I	9	19.5	19.5	0.00	19.5	19.5	0.00	
9	Geometry	10	24.5	23.5	0.06	28.0	20.0	4.80	*
Pooled			164.0	169.0	0.23	177.5	155.5	4.36	*

R_c: rank sum for comment group; R_{nc}: rank sum for no-comment group. Higher rank sum is better performance.

X_1^2: non-parametric two-way analysis of variance with one degree of freedom; * → significant at $\alpha = .05$; ** → significant at $\alpha = .01$.

4. Findings

The null hypothesis was rejected ($\alpha = .05$) in one geometry and one general mathematics class, with the comment group surpassing the no-comment group.

5. Interpretations

This study found support for increased achievement for the comment group in two of nine classes. These two were one geometry class and one general mathematics class. Recognizing the dangers of generalizing homework comments to test comments, this study does not appear to support either Page's results across all grade levels, or the results of Schoen and Kreye, who found no positive effect.

The time limit (six weeks) and allowing students to ask questions on problems in class may have mitigated the results of the comments on achievement. Finally, more research on effects of homework comments is suggested by the investigator.

Critical Commentary

The study was well designed and focused on a subject that has a practical meaning for the classroom and a theoretical meaning for mathematics education research. The findings suggest improvement in achievement

in certain subject areas, geometry and general mathematics, as a result of comments on homework papers, for two of the nine classes. It may be that the kind of knowledge studied (for example, concepts vs. skills) is a greater factor in dealing with comments on homework. This may be why students improved more on geometric proof than on algebraic skills.

A replication of the study with a larger sample of inservice teachers and an increase in the time of the experiment might be fruitful. The fact that no implication could be agreed upon concerning the fourth-grade sample may warrant further separate investigation at the elementary-school level on the effects of comments on homework papers on achievement.

This abstractor agrees with the author that more research is needed to investigate the effects of written comments on homework.

Boyd D. Holtan
West Virginia University

Page, E. B. Teacher Comments and Student Performance: A Seventy-Four Classroom Experiment in School Motivation. Journal of Educational Psychology, v49, pp173-181, 1958.

Schoen, Harold L.; Kreye, Bettibel C. Five Forms of Written Feedback to Homework in a Mathematics Course for Elementary Teachers. Journal for Research in Mathematics Education, v5, pp140-145, 1974.

THE CONTENT OF HIGH SCHOOL COMPUTER MATHEMATICS COURSES. Buchman, Aaron L.; Hawthorne, Frank S. School Science and Mathematics, v75 n6, pp543-549; October 1975.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by Jane Donnelly Gawronski, San Diego County Department of Education.

1. Purpose

This survey was designed to determine the content of computer mathematics courses in grades 6-12 in the schools in New York State.

2. Rationale

Courses in computer mathematics at the high school level are being offered in an increasing number of high school mathematics departments. However, there is a question about the actual content of these courses, since there are few textbooks available to serve as a major text and/or a resource for such a course. Mathematics teachers themselves have had to develop the course content. One would expect that after a teacher has taught a computer mathematics course for several years that a "substantial reservoir of apt, pertinent, interesting and challenging concepts and procedures" would have been developed. The Mathematics Education Bureau of the New York State Education Department believed this was occurring in New York State. An ad hoc computer mathematics committee called together by the Department in 1962 had recommended that no State course of study be issued in the interest of encouraging experimentation by teachers with course content. Thus, schools had been encouraged to develop their own computer mathematics course outlines for several years.

3. Research Design and Procedure

The Bureau prepared a questionnaire that contained a checklist involving 32 topics related to computer mathematics. The previous year the schools which offered a computer mathematics course had been requested to submit their locally developed course outlines to the Bureau. Topics in the questionnaire were abstracted from the combined contents of 28 course outlines submitted by the schools. For each topic in the questionnaire the teacher was asked to indicate:

(1) whether the topic was being taught at the school; and

(2) how the teacher rated it in relative importance using a five-point scale.

The questionnaire was sent to 173 schools during the school year 1971-72. Replies were received from 165 teachers in 138 schools, representing an 80% return from the schools canvassed.

4. Findings

The results indicated that 18 of the 32 topics were being taught in at least 60% of the schools. Looping and branching concepts, flow chart applications, iterative techniques and nesting procedures were computer topics near the top in usage. Use of the quadratic formula, solving systems of equations, prime number problems, and construction of tables and arrays are some of the other mathematics topics that appear high on the list. The remaining 14 topics were taught in from 15% to 55% of the schools surveyed. A total of 30 additional topics were "written in". Nine schools added "playing games".

Teachers also ranked the topics according to belief in their importance. The statewide correlation between rankings by practice and by belief was 0.92. The relationship between the number of years of experience teaching computer mathematics which the teacher and his or her rating of the listed topics was investigated. Teachers were grouped into three categories:

Those who had taught computer mathematics for:

(1) 2 years or less,

(2) 3 or 4 years,

(3) 5 or more years.

There were 54, 59, and 52 in each group respectively. The correlations among the subgroups and between the whole state and its subgroups were all very high (from .83 to .97).

Subgroups determined by geographical areas of the State were also examined. By happenstance, these areas also differed in their use of hardware. Use of time-sharing BASIC systems, use of stand-alone computers using FORTRAN, and use of programmable desk calculators were the differences identified. The resulting correlations were somewhat lower -- from .62 to .91.

5. Interpretations

Teachers who have experience teaching a high school computer mathematics course do seem to develop a large common core of topics which they teach and which they believe to be important. In addition, hardware has a strong influence on course content, on how the course is taught, on what topics are emphasized, and on what the teacher feels is important. In any case, teachers and administrators who are considering implementation of a computer mathematics course might consider the topics identified by this questionnaire.

Critical Commentary

This questionnaire study identified the state of the art for computer mathematics courses in New York State by identifying the topics taught. In addition, it managed to do a reality check by asking teachers to rank their

beliefs in the importance of the topics. Predictably, this correlation between rank of topic and rank of belief in importance was high. One can't help wondering if this was due to the fact that teachers have a great deal more autonomy over content in computer mathematics courses than in a mathematics course with a text or a State course of study to cover.

Computer mathematics seems by its nature to be hardware-bound, yet there were similarities in the three different geographical areas with emphasis on three different hardware configurations. Some computer-related topics can be taught without access to any hardware and this point should be noted.

It would have been informative to describe some of the teaching/learning conditions. The topics taught and hardware available is interesting but how are they taught and what are the characteristics of the students who elect and subsequently succeed or fail at computer mathematics concepts? What are the characteristics of teachers who seek to teach a course where texts and State courses of study are not readily available?

It was interesting that the teacher group with 3 or 4 years experience seemed to have slightly less resemblance to the others. The authors conjecture that this indicates some greater experimentation with new emphases by this group, to be followed with further experimentation and then followed by a return to the original thinking. An alternate conjecture is that peer pressure to conform forces this group to return to the original thinking. It is unfortunate, but one concern this raises is the thought that once a content area appears in a text or in a list of suggested topics it seems to acquire an apparent God-given right to be there. Change can be a slow and difficult process.

This study does identify what is being taught in computer mathematics courses in New York State. The topics might well serve as the starting place for one interested in implementing a computer mathematics course.

Jane Donnelly Gawronski
San Diego County Department
of Education

SYSTEMATIC ERRORS IN THE FOUR VERTICAL ALGORITHMS IN NORMAL AND HANDICAPPED POPULATIONS. Cox, L. S. Journal for Research in Mathematics Education, v6 n4, pp202-220, November 1975.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by Richard A. Lesh, Northwestern University.

1. Purpose

The purpose of the study was to identify the most frequently occurring systematic errors in vertical algorithms for addition, subtraction, multiplication, and division of whole numbers. Data were gathered from children in regular classrooms and in special education classrooms. To investigate whether systematic errors persisted over time, many of the children who made systematic errors in the initial testing session were retested after one year.

2. Rationale

The paper summarized the results of a comprehensive review of the literature concerning the diagnosis, remediation, and analysis of systematic errors in the four basic algorithms. From 1900 to 1973, research was found to be extremely limited concerning systematic errors in computational processes.

3. Procedure

For each of the four basic algorithms, problems at several levels of difficulty were investigated. For example, for the addition algorithm, eight problem types were identified representing skill levels ranging from the addition of a two-digit and a one-digit number (without renaming) to the addition of a three-digit and a two-digit number (with renaming) or the addition of three two-digit numbers (without renaming). Similarly, for the other three operations, several "skill levels" of problems were devised according to the number of digits that were involved and the inclusion of renaming. Altogether, six skill levels were identified for subtraction, ten for multiplication, and ten for division. For each skill level of each operation, five problems were posed. Responses were classified into one of the following categories:

(1) Systematic Error. The child missed at least three out of five problems because of a repeatedly occurring process of thinking.

(2) Random Error. The child missed at least three out of five problems of a particular type, but a pattern of error or a systematic incorrect process could not be detected.

(3) Careless Error. The child missed only one or two problems out of five of a particular type. He basically knew how to work the problem.

(4) No Error. All five problems were correct.

(5) Incomplete Data Sheet. The child did not work all five problems. An effort was made to keep this category as small as possible.

The study involved 744 children from several public, private, and parochial schools in the greater metropolitan Kansas City area, with approximately 115 children from regular classrooms in grades 2-6, and 180 children from special education classrooms in the above schools or from the University of Kansas Medical Center. Among the special education youngsters, 45 were at the primary level, 70 were at the intermediate level, and 65 were at the junior high school level.

Two criteria had to be met before a child's responses were analyzed for a particular type of problem: (a) the child had received prior instruction pertaining to the particular type of problem being studied, and (b) the child had an adequate knowledge of the basic facts relevant to the problem being studied.

4. Findings

Among the 744 children who were involved in the study, 223 systematic errors were identified. These errors were classified according to dysfunction and are described in detail in a major report (Cox, 1973). For example, for the addition algorithm, the 51 different systematic errors were grouped into four categories--renaming, concept of addition, wrong operation, and place value. The criteria for forming a grouping category was the hypothesis that the errors were so closely related that they could possibly be corrected using a single training procedure.

For each algorithm, the percentage of children who made systematic errors at each grade level was reported, as well as the percentage of systematic errors that were classified under each type of dysfunction. The results showed that more systematic errors were made on the subtraction algorithm (i.e., 13%) than on addition (5%), multiplication (6%) or division (6%). The data also showed that the occurrence of systematic errors was much higher in grades 2 and 3 than in grades 4, 5, and 6, and that, except for the junior high school special education students, the occurrence of systematic errors was much higher in special education classrooms. For example, in the third grade, systematic errors in subtraction occurred in 23% of the cases; and in intermediate-level special education classes, systematic errors occurred in 8% of the cases for addition, in 24% of the cases for subtraction, in 27% of the cases for multiplication, and in 23% of the cases for division. On the other hand, in regular sixth-grade classes, systematic errors reached a low of 0% for addition, 6% for subtraction, 1% for multiplication, and 3% for division.

To investigate whether systematic errors are only a transient problem or if they persist over time, certain skill levels in subtraction and multiplication were selected for analysis in a following study. Among the children who had made systematic errors in these categories, 191 children were selected to be retested after one year. In this sample, 23% were making systematic errors on the same algorithm one year later, and 59% of these were making the identical error as in the initial testing session.

5. Interpretations

Because the children in this study were screened for "knowledge of basic facts" and "prior instruction on the relevant algorithms," the high frequency of errors cannot simply be attributed to a lack of knowledge of basic facts. In fact, some of the errors which were classified as "random errors" were similar to systematic errors but did not occur with enough frequency to be defined as such. So, systematic errors could be even more of a problem than the data in this study indicate. Furthermore, unless something is done to correct the errors, they may persist over relatively long periods of time.

Critical Commentary

Cox's study is a commendable first step in the isolation and analysis of systematic errors in computational algorithms. Further studies of this type are clearly needed. Future studies might attempt to determine how errors on one algorithm are related to errors on another algorithm, or how errors at one skill level are related to errors at another skill level for the same algorithm. It would also be interesting to have a more thorough description of the abilities (or disabilities) of the children being tested. For example, in Cox's study the definition of "special education" children was rather "loose", involving "learning problems" classrooms as well as educable mentally retarded classrooms. Pupils were assigned to learning problem classrooms if they were having difficulty making progress in normal classrooms, and if it was judged they might benefit from special classroom placement. Nonetheless, many special education pupils, or learning disabilities children, have been given extensive series of diagnostic tests to pinpoint (insofar as possible) specific disabilities. Yet, little is known about the relationship between these underlying disabilities and skills on computation tasks like the ones in Cox's study. It seems likely that progress will only be made on these types of questions if large quantities of data are analyzed and reanalyzed by a number of people from a variety of theoretical perspectives. Consequently, data collection should be conducted keeping in mind the likelihood of future analyses.

Finally, it is of great importance to find instructional experiences that can effectively deal with problems of the type described in Cox's study. In fact, Cox herself has begun work in this area. When a class of systematic errors were classified as "place value" errors, the assumption was made that all of these errors could be corrected using a single instructional sequence whose focus was on place value. Yet, little research has been done to investigate the validity of this assumption or similar assumptions.

Richard A. Lesh
Northwestern University

THE OBSERVATION-INTERVIEW METHOD AND SOME CASE STUDIES. Erlwanger, Stanley. Proceedings of the Conference on the Future of Mathematical Education, pp125-142. Florida State University, Project for the Mathematical Development of Children. 1975.

Also reported in:

CASE STUDIES OF CHILDREN'S CONCEPTIONS OF MATHEMATICS - PART I. Erlwanger, Stanley Herbert. Journal of Children's Mathematical Behavior, v6 n2, pp157-283, Summer 1975.

Also see:

Dissertation Abstracts International 35A: 7121, May 1975. (Order No. 75-11,653)

Expanded Abstract and Analysis Prepared Especially for I.M.E. by Vincent J. Glennon, The University of Connecticut.

1. Purpose

Using the observation-interview method, to study and report on several interacting variables in the holistic teaching-learning situation of a group of fourth-, fifth- and sixth-grade children studying Individually Prescribed Instruction (I.P.I.) Mathematics for several years.

2. Rationale

The overwhelming majority of research studies in mathematics education reported in well refereed journals use nomothetically-derived data within a design that attempts to meet the two criteria of scientific research - repeatability of the study and the constancy of the data. Aware that certain learnings such as the child's "views about mathematics" and the processing of cognitive material are not adequately gathered by such formal methods of research, the author chose to use the more clinically oriented observation-interview (idiographical) method.

3. Research Design and Procedure

The researcher chose to use an unstructured ("informal, flexible and exploratory"), hence non-repeatable interview technique in which his next question, in a sequence of questions, was determined by the subject's response to the prior question. The interviewer attempted to establish a client-centered, non-threatening atmosphere in which the child could respond freely, and at times with considerable affect, to largely verbal-cognitive situations created by the investigator.

Partial interviews with two subjects are reported. All interviews were recorded on audiotape; each interview lasted about thirty minutes. The researcher identified four characteristics of his research design.

1. The children were in a mathematics program whose distinguishing features are its form of individualization and its carefully sequenced curriculum, with a corresponding testing program for monitoring a child's progress.

2. The children were above average in mathematics achievement according to their performance on tests, progress in the program, and the assessment of their teachers.

3. The children generally appeared to be independent, confident, and content with their progress in mathematics.

4. The children also revealed unanticipated underlying conceptions of mathematics and learning mathematics in some areas of their work.

He then raised the educationally significant question of reconciling observation 4 in the light of statements 1, 2, and 3.

4. Findings

Case Study 1: Knowledge of Rules and Answers

Benny is a twelve-year-old, sixth-grade boy. He has been in I.P.I. mathematics since the second grade and has been "making good progress." Hence, in the judgement of his teacher, Benny must have "mastered" his previous work in the program. The researcher, however, focused his questions on Benny's conceptual (rather than his associational) mastery of fractions and decimals for the purpose of identifying the child's ideas about answers and rules.

In so doing, the investigator uncovered a tangled web of incorrect, self-taught algorithms and rules for converting fractions to decimals, for addition and multiplication of decimals, and for addition of fractions. "...Through using I.P.I., learning mathematics has become a 'wild goose chase' in which he (Benny) is chasing particular answers." Benny believes that "In fractions we have 100 different kinds of rules."

The researcher further concluded that "through an emphasis on answers in the testing program, the (answer) key seems to associate his teacher with his frustrations. It appears, then, that in I.P.I. the teacher may be prevented by her role perception from understanding the pupil's conception of what he is doing."

Case Study 2: Knowledge of Methods and Answers

In working with Mat, an eleven-year-old boy with an IQ of 121, the researcher directed his interviews toward the child's conception of methods and answers in mathematics. Again, he found the child had through self-teaching invented, practiced and consolidated a large number of incorrect ideas related to the algorithms for processing whole numbers and fractions.

For instance:

To show 3/4, he divided a circle into four equal parts. Then he said, "You shade in 3 parts. Take away 3 parts ... 3/4." He also arranged four wooden blocks, and then removed three blocks and said: "I had 4 of them, and I take away 3 of them." This concept of forming a fraction a/b as take a from b led him to the following conclusion about 3/4 + 1/4:

(a) With wooden blocks, 3/4 + 1/4 = 4/4 = 0, because "Here you have 4 ... [and] here you are taking 4 out of 4 [blocks]."

(b) Using a circle diagram, 3/4 + 1/4 = 1: "If I divide that [circle] into 4, and then added that one [shading 3/4] and this one [shading 1/4] ... it's a whole ... [and] ... You have a circle with 4 in it, and you have all 4 ... that's a whole."

(c) Using his "rule", he said: "One way to do that ... 3 + 1 is 4, 4 + 4 = 8 ... 4/8 ... [and] ...2/4 or 1/2." He wrote: 3/4 + 1/4 = 1/2.

These (conflicting) results did not appear to disturb Mat. When he was asked: "How would you decide which answer to take ... or would you just accept all of them?" He said: "Well, I'd probably do the one where you add the denominators and the numerators, because that's what method they've (I.P.I.) taught me to do in the booklets ... and they didn't teach me to do it with diagrams."

With both children, Benny and Mat, the researcher pursued programs of remediation. In each instance he found that the incorrect concepts of rules and methods had been so heavily consolidated that they were resistant to change. New learnings were acquired slowly and with great effort.

5. Interpretations

The clinical approach to teaching and to diagnosis, as has been known and demonstrated for many years, has been once again to be a powerful professional technique for studying complex interacting variables in the teaching-learning process when working with individual children. The objectivity and reliability obtained in the scientific studies following the behavioristic methodology are sacrificed here to attain equally important goals of the neo-humanistic tradition. Each methodology is necessary; neither is sufficient.

The researcher demonstrated that much incorrect mathematical learning took place and that this learning is not easily remediated.

Critical Commentary

Erlwanger has made a substantial contribution to raising the awareness level of the community of mathematics educators toward the value of the clinical method. The two detailed verbatim accounts show how children, seemingly mastering the I.P.I.-Mathematics program, not only failed to learn correct mathematics but, even more tragic, learned and consolidated much incorrect mathematics. Indirectly, he also raises serious questions about this and similar mathematics programs in which the role of the

teacher is shifted from regular, structured teaching to conceptual-developmental lessons in which teacher and children interact, to a program in which each child works alone and in which most of the teacher's time is devoted to managing the system rather than instruction.

Three questions occur to this abstracter:

1. How much additional idiographic data would have (should have) been obtained by using videotape rather than audiotape? Much useful non-verbal communication by the two children was undoubtedly lost.

2. How much idiographic data on, say, attentional dysfunction, the anxiety level raised by the cognitive dissonance situations, and other affective variables were present but not collected?

3. If the investigator had used one of the readily available models for diagnosis would he have bridged somewhat the gap between the behavioristic and humanistic schools and thereby increased the validity of the clinical method?

Vincent J. Glennon
The University of Connecticut

THE EFFECT OF SCHOOL ATTENDANCE ON INTELLECTUAL DEVELOPMENT IN NORTHERN NIGERIA. Fahrmeier, Edward D. Child Development, v46 n1, pp281-285, March 1975.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by Lewis R. Aiken, Sacred Heart College, Belmont, North Carolina.

1. Purpose

The question of whether schooling has an effect on cognitive development was investigated with predominantly Hausa children in a Northern Nigerian town of approximately 16,000 people. The primary hypothesis was that performance on selected culture-specific and "Western" tests would improve with age, with performance improving at a faster rate in school attendees than in non-attendees. Test differences between attendees ("schooled") and non-attendees ("unschooled") should be greater for older than for younger children, since the latter were just beginning school. The smallest performance differences between schooled and unschooled children were expected on orally administered tests such as vocabulary and information, which are highly loaded with culturally familiar material. It was also acknowledged that the effects of schooling would not appear on tests that were too difficult or measured skills not developed in late childhood or elementary school.

2. Rationale

Reference is made to previous investigations that compared schooled and unschooled children in cultures having limited educational facilities. When only a few tasks are employed, these studies have found either major differences or no differences between schooled and unschooled groups. Mixed results are obtained when a greater variety of tasks are used. The author proposed the present study to serve as a data base for conducting a thorough analysis of the particular culture.

3. Research Design and Procedure

The sample consisted of all children aged 6-13 who lived in the central ward of the selected Northern Nigerian town. The schooled sample consisted of 230 children, almost all of whom were located and tested. Only 175 of the 270 unschooled children, who were harder to locate and less willing to be tested, were given any part of the test battery. Although only 30 of the unschooled children took the entire battery, at least 70 test protocols were obtained for each test over the entire age range. The exact ages of 100 children were obtained from the town birth register. The ages of the remaining children were determined from a combination of fathers' estimates and ages predicted from height-age correlation computed on the known sample of 100.

Six culture-specific tests (1. Vocabulary, 2. Digit Span, 3. Memory of Verbs, 4. Memory of Objects, 5. Mental Arithmetic, 6. Information) and

six "Western" tests (7. Raven's Colored Progressive Matrices, 8. Cylinder Matrix, 9. Goldstein-Scheerer Cube Test, 10. Figure Grouping, 11. Spatial Relations, 12. Formboard) were administered to children aged 9-13. Only tests 1-8 and 12 were administered to children under age 9. Four male Nigerian secondary school graduates who spoke fluent Hausa administered the tests. Tests were given on porches, in entrance rooms, and under trees, with testing conditions being similar for schooled and unschooled children. The administration sequence was tests 7 and 8, 1, 2 through 6, and 9 through 12. Almost all children were tested within a two-month period.

4. Findings

Scores on each test were analyzed statistically by analysis of variance, the factors being age (6-7, 8-9, 10-11, 12-13), schooling (schooled, unschooled), and sex (male, female). The main effect of age was significant in favor of the older children on all tests except Spatial Relations, Figure Grouping, and Memory of Objects. Significant age differences in test scores were accounted for by the linear components in trend analyses, with no quadratic trend being significant. The main effect of sex was significant only in Figure Grouping, the Cube Test, and Memory of Objects. Girls did better on Memory of Objects and boys on Figure Grouping and the Cube Test. The anticipated interaction between the schooling and age variables did not materialize. Neither did the relatively higher scores that the schooled children might be expected to make on the Western-style than on the more culture-specific tests.

The overall results are summarized as indicating that age and schooling were related to higher test scores, but the differences in test scores between schooled and unschooled children tended to be constant across age levels.

5. Interpretations

The investigator considers several interpretations of the finding that the mean test score differences between the schooled and unschooled children remained fairly constant across age levels in favor of the former group. One possibility is that the children who are sent to school are more alert and attentive than those not sent. Support for this explanation is found in the fact that the unschooled children appear to be at a disadvantage as early as age six. But the rate of improvement in the unschooled children is parallel to that of the schooled children, and a more rapid improvement in the latter would be expected if the age differences between the two groups were due to differences in native intelligence. The investigator accepts as a more likely explanation that the constant difference between the groups is due to differences in motivation to do well on school-type tasks. Whatever the correct explanation of the findings, it seems clear that the skills measured by the tests used in this study are also learned out of school. The investigator concludes that attending school classes, in which topics are handled more abstractly, does not appear to have a major effect on cognitive growth.

Critical Commentary

This study is another demonstration of the difficulty of interpreting results and establishing causal connections from non-experimental investigations. As such, it is representative of much educational research, not only that conducted in non-Western cultures but in our own culture as well. Without adequate controls, one simply does not know what organismic or environmental variables produced the results obtained here. Other methodological difficulties are: (1) the very small sample sizes of some subgroups, especially in the unschooled children; (2) problems in determining the exact ages of many of the children; (3) inappropriate difficulty level of some of the tests; and (4) questionable conditions of test administration.

With the exception of Spatial Relations, the three tests that are most related to mathematics education -- Digit Span, Arithmetic, and Spatial Relations -- behaved similarly to the other nine tests. It is curious that there was no sex difference on the Spatial Relations test, since American boys usually score better than American girls on this type of test. However, the sex differences in favor of boys on the Figure Grouping and Cube Test and in favor of girls on the Memory for Objects test are consistent with other research findings.

Future non-experimental investigations of this sort should probably attempt to employ some kind of matching procedure, or at least concurrently measure other variables such as achievement motivation that may serve as alternative explanations for the findings. Furthermore, it cannot be concluded from the data obtained in this investigation that schooling does not affect cognitive development. A more plausible conclusion is that the tests and methodology used did not do an effective job of assessing whatever differences schooling makes.

Lewis R. Aiken
Sacred Heart College
Belmont, North Carolina

LONGITUDINAL INVESTIGATION OF TEACHER-STUDENT INTERACTIONS AND THEIR RELATIONSHIP TO ACADEMIC PERFORMANCE. Firestone, Glenn; Brody, Nathan. Journal of Educational Psychology, v67 n4, pp544-550, August 1975.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by John W. Gregory, University of Florida.

1. Purpose

This longitudinal investigation was performed in order to determine the relationship between affective qualities of teacher verbal responses and the achievement of first-grade students.

2. Rationale

Several investigations are cited which lend credence to the hypothesis that the quality of interactions between a teacher and a student expresses differential teacher expectation and in turn is related to academic achievement. Rosenthal and Jacobson's controversial work Pygmalion in the Classroom is one example of a study in the realm of expectation-achievement research. Other studies cited have investigated the relationship between teacher expectation and the affective quality of teacher-student interactions. The study to be reviewed sought the third-pair relationship (affective interaction-achievement).

3. Research Design and Procedure

Sixty-three children from four different classes in one school and their six kindergarten and/or first-grade teachers comprised the sample. The children were non-white with approximately 57% from families eligible for welfare.

Over the period of one year covering the last half of kindergarten and the first half of first grade, a two-hour observation was made in each class every three weeks using Flanders Interaction Analysis System. In addition to the quality of teacher interaction offered by this five-second interval system, the number of times a child was chosen to demonstrate something for the class and to whom the teacher verbal behavior was directed were also recorded. A reliability check was offered on one observation by another trained coder for which 89% agreement was established. For these observations, neither content nor group size involved in the interaction with the teacher was held constant between classes or observations.

The observations led to the computation of the following interaction scores for each child:

a. percentage of positive interaction (teacher accepts feelings, praises or encourages, accepts or uses ideas) for each quarter of the year during which observation data were obtained.

b. percentage of negative interaction (teacher criticizing or justifying authority) for each quarter.

c. total frequency of interactions (positive, negative and neutral) for each quarter.

d. total frequency of times asked to demonstrate.

Dependent measures were derivable from performance on the Metropolitan Achievement Test administered by the first-grade teachers at the end of the year. Five subscores of Word Knowledge, Word Analysis, Reading, Total Reading, and Total Math constituted five additional dependent measures. Reliabilities are not reported.

Multiple regression analyses was used to predict the MAT scores on the basis of IQ, Primary Academic Sentiment Scale scores (administered with the MAT; description not provided), sex, interaction scores, and welfare status of the child's family (not defined).

4. Findings

1. The percentage of negative interaction in the second quarter showed a significant and consistent negative relationship to scores on all subtests of the MAT.

2. The total "number of times asked to demonstrate for the year" related positively and significantly to scores on Word Knowledge and Total Reading subtest scores.

3. Total number of interactions in the second quarter related positively and significantly to Total Math subtest scores.

4. The percentage of positive interaction for the third quarter related significantly and negatively to scores on Word Knowledge, Word Analysis, and Total Reading subtests.

5. When IQ was deleted from the analysis, the percentage of negative interaction emerged as a significant source of variance (with a negative relationship).

6. With IQ deleted as a source of variance, the "number of times asked to demonstrate" became a significant and positive source of variance in all subtests.

7. Welfare status entered as a significant positive correlate of scores on all subtests when IQ was deleted.

5. Interpretations

1. The emergence of significant variables when IQ variance is deleted indicates the dependency between those variables and IQ.

2. Children who experience the highest percentage of negative interaction with their kindergarten teachers did more poorly on the MAT both with and without accounting for IQ. The investigators felt that this implied that negative interaction does not relate negatively to achievement, solely on the basis that students being "duller" results in more negative interactions.

3. One explanation offered relates to the self-fulfilling prophecy: being treated negatively leads children to believe they are not worth much; not much is expected of them and therefore they try less hard and achieve only to the level of expectation. Similarly, those asked to demonstrate more frequently get an ego boost and try harder.

4. Another explanation involves a third variable possibly being the cause of behavior of the teacher in relation to particular students (like the personality or temperament of the student).

Critical Commentary

As another investigation which adds to the body of knowledge relative to the relationship between particular verbal interaction qualities of the classroom environment and student product and process variables, this study is important. Questions which may direct future efforts of this type are listed below.

1. Are there some topics which result in more negative interaction for particular teachers and/or students? And if so, do variables such as teacher (student) attitude or ability in that topical area provide a more basic, underlying motive for the negative interaction? Holding content constant for all observations would at least control for such possibilities.

2. Although means and standard deviations for scores relative to the interaction are not reported, does reluctance on the part of students to become involved in interaction due to shyness or mere indifference cloud data of this type? Could it be assumed that all children are affected by negative interaction regardless of to whom the interaction is presented and that the total verbal environment rather than personal instances of positive and negative interactions provide a better measure? Thus teachers and entire classes serving as the sampling unit would be compared rather than individual students.

3. If part of the theory indicates that interactions are related to motivation, might not the scores on the Academic Sentiment Scale be best used as a dependent variable rather than as an independent variable?

4. If negative and positive interactions provide some causal influence on achievement (not determinable from this study), is there an optimal frequency or percentage? Both interaction

types of negative and positive responses led to negative correlates for various aspects of achievement. Might this not imply that it is possible that a teacher can use too much positive interaction as well as too much negative interaction? Is there an optimal balance of the two types?

John W. Gregory
University of Florida

FOLLOW-UP OF HIGH-RISK CHILDREN IDENTIFIED IN KINDERGARTEN THROUGH DIRECT CLASSROOM OBSERVATION. Forness, Steven H.; Guthrie, Donald; Hall, Robert J. Psychology in the Schools, v13 n1, pp45-49, January 1976.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by Norman L. Webb, University of Wisconsin-Madison.

1. Purpose

To determine the validity of clusters of observable behavior of beginning kindergarten children in predicting the educational risk of these children at the completion of first grade.

2. Rationale

Direct classroom observation of children is one possible means of providing information to predict later school progress. This is the third of three studies using data gathered by direct observations on the same population of four kindergarten classes. The first study found that the observational data gathered at the beginning of the school year predicted children who were identified by their teachers toward the end of kindergarten as educationally high-risk. The second study used cluster analysis and revealed that high-risk children tended to be less attentive, to be more active in both on- and off-task situations, and to require more teacher supervision than non-risk children in the same classroom. This study is a follow-up of the original kindergarten population to test the validity of the original clusters of observable behavior in predicting educational risk at the completion of first grade.

3. Research Design and Procedure

The population used for the first two studies were 94 children (51 boys and 43 girls) enrolled in four kindergarten classes of a racially mixed public elementary school in a lower middle class neighborhood in West Los Angeles. Sixty-one children of the 94 were still enrolled in the same school and were used as the population for the follow-up study. A chi-square test and analyses were used to judge that the student attrition was randomly distributed and that the follow-up group was relatively representative of the original population.

The original 94 children were divided into four clusters based on classroom observations at the beginning of kindergarten. The 61 remaining children in first grade were then compared by clusters on two measures of school progress (teacher ratings and achievement tests) to test the predictive validity of the clustering.

Children in each classroom were observed using a time-sampling procedure at six-second intervals in round-robin fashion. This was continued until each child had been observed a minimum of 10 intervals per day. Classes were observed on consecutive school days until all children in the population had been observed a minimum of 10 days (mean = 13.8 days).

The observable behaviors were recorded in four predetermined categories: Verbal Positive - task oriented verbal or gestural response; Attend - eye contact to teacher, task material, or peer who is reciting; Non Attend - nontask-oriented eye contact; and Disrupt - behavior incompatible with on-task activities.

A cluster analysis was performed on the percentages of total behavior in each category to total observed behavior. Four clusters were identified:

		Percentage of Observable Behaviors			
Cluster	N	Verbal Positive	Attend	Non Attend	Disrupt
1 nonrisk	25	Low (4.5%)	High (86.7%)	Low (0%)	Low (0%)
2 seen as nonrisk	24	Med. (6.9%)	Med. (72.9%)	High (20.2%)	Low (.03%)
3 seen as high risk	29	Med. *	Med. *	High *	Med. (1.3%)
4 high risk	16	High (8.8%)	Low (63.8%)	High (21.9%)	High (5.5%)

*Percentages not given.

These categories were essentially confirmed by concurrent teacher ratings.

Two measures of educational progress were collected for each child at the end of first grade:

a. Teacher evaluations in reading, arithmetic, classroom behavior, and relationships with other children using a five-point scale (1 - poor to 5 - good), and

b. Peabody Individual Achievement Tests (reading recognition, mathematics, spelling, and general information).

F-values were computed to assess the cluster effect and the predictive validity of clustering for teacher ratings and achievement scores, using as covariates initial teacher rating in (1) kindergarten and (2) kindergarten and first-grade classroom placements. Pearson correlation coefficients between each observation category and total first-grade achievement scores were computed.

4. Findings

Cluster 1 (nonrisk) had the highest mean scores in all four areas of the teacher ratings and in each of the four achievement measures. Cluster 3 was next highest in mean scores on the teacher ratings. There was no consistent order for clusters 2 and 4 on the mean scores of the teacher ratings or for clusters 2, 3, and 4 on the mean scores of the four achievement measures.

Some cluster effects, using initial teacher rating in kindergarten as a covariate, were found at the completion of first grade to be significant. These include total teacher ratings ($p < .05$), arithmetic teacher ratings ($p < .02$), total achievement scores ($p < .002$), reading subtest scores ($p < .05$), mathematics subtest scores ($p < .05$), and spelling subtest scores ($p < .05$). No significant cluster effects were found due to instructional placement in either year using kindergarten and first-grade classroom placement as covariates.

In regard to predictive validity, cluster 1 children were significantly higher in total achievement and total teacher rating. Using kindergarten teacher rating as a covariate, cluster 1 children were significantly higher only in total achievement; cluster 2 children were significantly lower in total achievement and on teacher ratings on arithmetic and reading; cluster 4 children were rated significantly lower in behavior and peer relationships.

Correlations between observation categories (VP, AT, NA and D) and total first-grade achievement scores were -.14, .42, -.45 and -.27, respectively.

5. Interpretations

The direct and systematic observations had relatively high inter-observer reliabilities (.87 to .89). The predictive validity of the observations only had partial success. Cluster 1 (nonrisk) children continued to do well at the end of first grade in both classroom behavior and academics. Children in cluster 4 (high-risk) were still doing poorly in most areas but only significant in classroom behavior. Cluster 2 children (seen as nonrisk) appeared to have dropped somewhat in academic standing. Achievement in specific content areas (reading and mathematics) was not predicted significantly for any one cluster.

Kindergarten teacher ratings were better than observational data in predicting ratings by first-grade teachers. This may be the result of "self-fulfilling prophecies" passed on to first-grade teachers. Teacher perception of educational risk may depend more on problem classroom behavior than on academic performance. Evidence of this is that cluster 4 children were lower in classroom deportment but equivalent in academics.

Attention may be a critical dimension in predicting educational risk. The level of classroom attending behavior of cluster 4 children was, in fact, equal to that of educationally handicapped children.

Training of classroom teachers in observational systems might insure that less conspicuous problems are more readily identified, supply data more directly related to a child's progress, and provide a functional analysis of behavior which pinpoints specific intervention strategies.

Critical Commentary

Great care must be taken in developing any instrument or technique to be used as predictors of future educational progress. Reliability and validity must be established. Once a satisfactory technique has been

developed, the important question still remains: What can be done to improve the educational experience of those "doomed" to failure?

The direct and systematic observations appeared to have intercoder reliability, but their predictive validity is still in question. Data from observations were used to identify the cluster of children who did, on the average, the best on achievement and teacher ratings. With the exception of this one cluster, the stratification of clusters did not occur.

Questions arise about the use of direct observations and whether or not there are more efficient and valid means of predicting future progress. No pretests were given, so the observation technique could not be compared with the use of readiness tests or other cognitive measures as predictors. Teacher ratings in kindergarten predicted fairly well teacher ratings at the end of first grade. The observations did provide some information beyond kindergarten teacher ratings, but only in a few cases. The real value of direct observations may be to suggest specific intervention, but this was not explored in this study.

There are questions as to whether the four clusters really differed on the predictive variables. The total time each student was observed was approximately 10 to 15 minutes. Is this enough time to characterize the student? Was this time spread out during the day so children were observed in all aspects of learning? What percentage of total classroom time was spent performing these four behaviors? Are there other behaviors that may be observable and more "telling", such as the number of questions asked or the time spent in individual work? Is 4.3% (the difference between low and high Verbal Positive behavior) really enough to distinguish between clusters?

Only mean performances were reported. In any clustering technique a wide variance within clusters can exist. How homogeneous were the clusters? What is the possibility of error in including a child in a cluster in which he or she does not belong or excluding a child from a cluster where he or she does belong?

There are values in the early identification provided that diagnostic information can also be obtained and corrective procedures suggested. Future attempts to validate the use of observations as predictors should take this into consideration.

Normal L. Webb
University of Wisconsin-Madison

THE EFFECTS ON PRESERVICE ELEMENTARY TEACHERS OF LEARNING MATHEMATICS AND MEANS OF TEACHING MATHEMATICS THROUGH THE ACTIVE MANIPULATION OF MATERIALS. Fuson, Karen. Journal for Research in Mathematics Education, v6 n1, pp51-63, January 1975.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by Kenneth A. Retzer, Illinois State University-Normal.

1. Purpose

The aim of this study was threefold:

(1) the development of sufficient curricular materials using manipulatives for a combination mathematics and mathematics education course,

(2) the creation or adaptation of ways of assessing learning from such materials, and

(3) an examination of some effects of such teacher learning.

2. Rationale

Citing (a) interest in manipulative materials in sessions of NCTM conventions and (b) short-duration studies on the effects of manipulative materials on achievement and attitude toward mathematics of college-age students, the researcher designed the study because few manipulative curricular materials for teachers then existed and research instruments for measuring results using them were largely nonexistent.

3. Research Design and Procedure

The treatment was a combination mathematics and mathematics methods course consisting of 20 two-and-one-half hour sessions at the University of Chicago in the fall quarter, 1971. An outline of topics covered and a partial list of manipulative materials used in the course is contained in this report. The topics, common to numerous courses of this type, included a study of numeration systems; the meanings of operations of the whole numbers and their implications for computational algorithms; set algebra; geometry; mathematical structure; measure; graphing as informational and problem-solving tools and as an introduction to functions and relations; the concepts and computational algorithms for integers, rationals and decimals. A synthesizing of topics summarized class experience, helped the students make and use manipulative materials, and informed them about models for mathematics learning or basic mathematics needed by each citizen. The manipulatives included poker chips, Cuisenaire rods, multibase arithmetic blocks, attribute blocks, geoboards, Mira, and a variety of common objects. Outside reading in pedagogy was discussed only briefly in class. A typical class session involved active manipulation of objects by trainees--individually or in small groups. The progression was from work with manipulative materials and recording of results to an analysis of the implication(s) of these results for the symbolic mathematical concepts involved.

The subjects were sixteen Master of Science in Teaching students who were college graduates with little or no prior training in education but who planned to become elementary teachers in a four-quarter program of study. Their preparation, ability, and self-image with respect to mathematics varied over a considerable range, but their IQ and verbal ability was higher than for preservice elementary teachers in general. The course was designed by the researcher and taught by a professor who was familiar with manipulative materials but had used them only briefly in prior teaching.

4. Questions and Findings

The effects of the treatment in this course were posed in five questions and are listed along with the findings as in the original report.

Question 1:

Does an extended learning experience that uses manipulative materials in a learner-focused manner (a) increase the desire of trainees to use manipulative materials in teaching mathematics, (b) increase the ability of trainees to use manipulative materials in teaching mathematics, and (c) increase trainee use of manipulative materials in teaching mathematics?

A post-course interview, rating of the teaching of eight mathematical concepts, and observing trainees while practice or student teaching in mathematics in the subsequent academic quarter were used to support the contentions that: trainees expressed an increased desire to use manipulative materials, they increased in their ability to demonstrate mathematical concepts with such materials, and they did use manipulative materials to a considerable extent in practice teaching.

Question 2:

Does an extended learning experience that uses manipulative materials in a learner-focused manner (a) increase the desire of trainees to teach mathematics in learner-focused ways and (b) increase the extent to which trainees teach in learner-focused ways?

Analyzing the results of use of a 20-item, five-point Likert scale instrument developed by Collier, which was given immediately before and after the course, indicated that the trainees expressed an increased desire to teach mathematics in learner-focused ways. Analyzing three-minute videotapes of each trainee partially teaching three concepts which were not explicitly covered in their course according to five rating categories indicated that they did increase the extent of teaching in learner-focused ways in four of five categories.

Question 3:

Does an extended learning experience that uses manipulative materials in a learner-focused manner increase trainee understanding of elementary mathematics?

Trainees self-reported an increase in understanding elementary mathematics concepts, and the researcher seemed to notice it although no precise measures of that increase exist.

Question 4:

Does an extended learning experience that uses manipulative materials in a learner-focused manner have effects on adults that differ from those of symbolic, abstract types of learning?

In trainees' descriptions of sudden insights, a questionnaire on mathematical and pedagogical learning administered five times during the course, and post-course interviews, trainees frequently used verbs directly connected with physical sensations such as "see" and "feel". This understanding in a physical and sensory way seemed to be the only distinction, since more symbolic and abstract materials can also engender sudden revelations of insight, enable students to explain a concept, and enable children to discover concepts for themselves.

Question 5:

Does an extended learning experience that uses manipulative materials in a learner-focused manner affect two aspects of trainee attitude toward teaching mathematics: enjoyment of teaching mathematics and a feeling of competence in teaching mathematics?

Trainees showed significant mean changes toward positive attitudes with respect to mathematics on an instrument which asked them to rate mathematics with four other subject matter areas on a five-point scale and also to rank these five areas with respect to trainee anticipation of enjoyment of teaching and feelings of competence in teaching those subjects.

5. Interpretations

The researcher outlined limitations of lack of a control group, use of only one instructor and subjects who were atypical of elementary teacher trainees, and cautioned concerning generalizability of results of this study pending further research.

Indicating that trainees had to translate experiences as students into behaviors appropriate for a teacher and that this learning experience did result in changes in trainee attitudes, abilities and teaching behavior, the researcher implies that certain kinds of experience as a student are sufficient for some to accomplish changes in their teaching behavior. This is the first of two implications seen for the training of teachers. The other implication is that teachers can learn mathematics in a way that is relevant to, and useful in, teaching mathematics to children. This is seen as an opportunity to provide more time for trainees to learn mathematics than by taking mathematics courses they perceive as unintelligible and useless.

Intense involvement and moments of great excitement, though unmeasured, are considered important effects of this study, and the trainees' descriptions

of this are taken to indicate that, for some persons, learning mathematics through active manipulation of materials helps provide a verbal superstructure of mathematical knowledge with concrete referents and seems to change trainee beliefs and behavior toward learner-focused use of manipulative materials in the teaching of mathematics.

Critical Commentary

Critique of the study methods and implications should be interpreted in light of the fact that these methods were a pioneering effort in studying the effects of such a treatment. The rationale for choosing the methods over alternatives was not reported and thus several questions arise. Why were the methods chosen to measure an increase in desire superior to direct questions posed by an impartial individual? Was there too massive an analysis for the size of teaching behaviors sampled to answer question 2? Would an achievement test rather than self-reporting better answer question 3 and provide a basis for comparison of learning with other treatments?

Contrasting this treatment with those which are more symbolic and abstract, it does not seem surprising that the only difference is the use of terms that describe learning in a physical and sensory way because that would seem to be implicit in using manipulative materials. It is encouraging that this method produces some of the same desirable effects as contrasting methods.

For teacher training we find that this learning situation enabled, and perhaps encouraged, teachers to teach as they had been taught. Perhaps this provides at least an equal amount of evidence of a modeling effect upon teaching as of evidence of the use of manipulatives in a learner-focused format.

The content and effects of the pedagogical readings outside of class are unreported, and in view of the general lack of research evidence of superior effectiveness of any one teaching method over others, one does not see in this research an indication that the training of teachers should move in the direction of emulating the treatment used in this study. However, the preparation of manipulatives and sequencing their use may be an extremely valuable contribution for those teachers or educators who choose to emulate the treatment of the study in their courses. Their teaching training programs can be affected by this information when it is disseminated.

As indicated, the study demonstrated that teachers can learn mathematics similar to ways children can learn it and in such a way that it is perceived to have more immediate payoff in classrooms. Since it captivates attention and increases involvement, it seems to be an exciting alternative to conventional mathematics and mathematics education courses for preservice elementary teachres. However, the comparison of actual knowledge of mathematics and comparison of knowledge of a variety of methods to teach this mathematics awaits the results of comparative studies. It does not appear to this reviewer that a theory of teaching is implicit in a theory of learning and, if that is so, neither can a variety of teaching methods be implicit in an extended learning experience of one particular type.

The researcher is to be commended for a timely pioneering effort which provides a basis for subsequent similar studies.

Kenneth A. Retzer
Illinois State University

A COMPARISON OF ACHIEVEMENT AND ATTITUDE OF STUDENTS USING CONVENTIONAL OR CALCULATOR-BASED ALGORITHMS FOR OPERATIONS ON POSITIVE RATIONAL NUMBERS IN NINTH-GRADE GENERAL MATHEMATICS. Gaslin, William L. Journal for Research in Mathematics Education, v6 n2, pp96-108, March 1975.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by Jane O. Swafford, Northern Michigan University.

1. Purpose

This study assessed the effects on achievement and attitude of a calculator-dependent algorithm set and the conventional algorithm set, the latter with and without calculators, for computing with positive rational numbers.

2. Rationale

Given the growing availability of handheld calculators and the persistent problem of identifying successful strategies for developing fractional computation skills with low-ability or low-achieving ninth-grade students, an investigation of the effects on achievement and attitude of an algorithm set which is calculator-dependent is timely. No research directly relating to this problem could be cited.

3. Research Design and Procedure

The subjects for the study were students who were randomly assigned to three ninth-grade general mathematics classes in each of two schools. The classes at School A were taught by the investigator. Three different treatments were used at both schools. They were:

T1 - Conventional algorithm set without calculators
(Operations on fractions were performed according to the "usual" textbook approach.)

T2 - Conventional algorithm set with calculators
(Same as T1 except students were allowed to use calculators throughout the unit and on all tests.)

T3 - Alternative algorithm set with calculators
(Operations on fractions were preformed by converting each fraction to a decimal on the calculator, then performing the indicated operation on decimals with the calculator.)

The instructional materials for all three treatments were developed by the investigator and used an individualized, mastery-learning model. The materials consisted of 5 units, one on each of the four basic operations with fractions and one on mixed operations. In addition, instruction in the use of the calculator was incorporated into the T2 and T3 instructional materials. The T3 group also received instruction in changing fractions to decimals.

Pretesting was done on prerequisite skills, operations on positive rationals, reading level, and IQ level. The first test worksheet completed in the fifth unit served as the posttest for the study. Each unit contained eight parallel test worksheets generated by the Honeywell Arithmetic Test Generator. A transfer-oriented posttest was also developed and administered. Students' attitudes were measured pre- and post-treatment using a 17-item semantic differential test. Following the post-treatment tests and a two-week unit on topology, a parallel form of the fraction pretest was administered as a retention test.

One- and two-way ANOVA and two-way ANCOVA were used to test six null hypotheses. Significant differences were further analyzed pairwise using the Newman-Keuls procedure. For the statistical analysis, each school was considered as a separate experiment, and the results were analyzed and reported separately.

4. Findings

a. There was a significant difference on the fraction posttest in School A favoring T3 over both T1 and T2. No significant difference was found among the three treatment groups in School B.

b. There was a significant difference in School A on 3 of the 6 parts of the transfer-oriented posttest, favoring T3 in all cases. In School B only one significant difference was found, favoring T2.

c. On the retention test a significant difference was found favoring T3 at both School A ($p < .01$) and School B ($p < .05$). The proportion of students who retained mastery level performance on the retention test at both schools favored T3, although these data were not analyzed.

d. There was no significant difference at either school among the three treatment groups on measures of attitude.

e. The mean number of days required by each treatment group at School A was reported. These data were not subjected to statistical analysis.

f. No significant interaction effects for reading ability level by treatment were found at either school on the transfer-oriented posttest, the fraction-retention test, or the attitude instrument.

5. Interpretations

From the findings, the investigator concluded that the alternative algorithm set with calculators appears to be a viable alternative for developing computational skills with fractions among low-ability or low-achieving ninth-grade students. While it probably should not be used with students who might eventually study algebra, it does provide a means by

which some students can compute with fractions. Furthermore, the alternative algorithm set was cited as producing successful transfer, as promoting superior retention, and as more efficient in terms of rate of mastery. The investigator further concluded that even with the conventional algorithm set, the calculator does not affect computational skills.

Critical Commentary

The study was well designed and carefully executed. However, the findings at the two schools are contradictory. Hence the conclusions would seem unjustified. Except for the retention test, the only significant differences favoring the alternative algorithm set with calculators were found in School A (in classes taught by the investigator). Even on the retention test, the differences at School B were not significant at the .01 level. Furthermore, the alternative algorithm set only produced successful transfer in the hands of the investigator. At School B, significant transfer was limited but favored a conventional algorithm group. In terms of rate of mastery, no data were available from School B. At School A, the alternative algorithm set appears more efficient only when the calculator related instruction days are unwarrantedly subtracted from the the mean number of days required to complete the unit. Lastly, while the investigator raised the issue of machine-dependent learners, his study did not address it. The conclusion that the use of a calculator does not affect computational skill was based on posttest results for which the calculator groups used calculators. Whether or not these students' computational skill is machine-dependent remains speculative.

The investigator established that he could successfully use an alternative, calculator-based algorithm set with low-ability students. But one wonders whether any significant differences would have been found had the data from the two schools been pooled. Generalizations about the success of the alternative algorithm set for low-ability or low-achieving students based on this study must be made with caution.

Jane O. Swafford
Northern Michigan University

LEARNING ENVIRONMENTS, ACADEMIC SELF-CONCEPTS, AND ACHIEVEMENT IN MATHEMATICS. House, Peggy A. Journal for Research in Mathematics Education, v6 n4, pp244-252, November 1975.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by Frank K. Lester, Jr., Indiana University.

1. Purpose

This study sought to determine whether or not the learning environment is a satisfactory predictor of academic self-concept.

2. Rationale

Three research studies indicated that students' academic self-concepts are related significantly to school achievement and that self-concepts improved as achievement increased. Another group of research studies indicated that certain characteristics of the classroom learning environment could be used as predictors of academic achievement. Since self-concept of academic ability and school achievement are related significantly and since characteristics of the learning environment can be used to predict achievement, the question arises whether the learning environment can be used to predict self-concept. If such a relationship could be shown to exist, further research would be warranted to determine the effect of classroom climate on academic self-concept and subsequent learning outcomes.

3. Research Design and Procedure

The researcher hypothesized that a student's perception of the classroom climate contributes to his or her academic self-concept, thereby affecting achievement indirectly through the student's self-concept. Self-concept of academic ability is defined as "behavior in which one indicates to oneself, publicly or privately, one's ability to achieve in academic tasks as compared with others engaged in the same task" (p. 245). General self-concept of academic ability (SCA) was assessed by means of an instrument developed by another researcher. A different form of this instrument measured self-concept of ability in mathematics (SCAM). Students' perceptions of the classroom environment were measured by means of a Learning Environment Inventory (LEI). The LEI was composed of fifteen scales: cohesiveness, friction, and competitiveness (measuring pupil-pupil relationships); formality, favoritism, disorganization, and democracy (measuring pupil-teacher relationships); diversity, speed, satisfaction, and difficulty (measuring pupil-subject matter-learning method relationships); and environment, goal direction, cliquishness, and apathy (measuring structural characteristics of the class).

Subjects were 1,079 ninth-grade algebra and non-algebra mathematics students enrolled in 15 junior and senior high schools. The schools ranged in size from under 200 to over 1,500 students and represented different organizational patterns with respect to the number and level of grades in the school. Five student characteristics were recorded as dichotomous

variables for each subject: course taken (algebra or non-algebra), school level (junior or senior high school), school size (small or large), sex, and whether or not a student had her or his present mathematics teacher in any previous course.

Multiple regression analysis was used to test the general hypothesis that no scale or combination of scales from the LEI could significantly predict either SCA or SCAM scores of subjects when one or more of the dichotomous variables were controlled. Five multiple regression analyses were performed, four controlling course, school level, school size, and sex; the fifth simultaneously controlled course, school level, and school size. In each analysis an attempt was made to predict SCA or SCAM scores using as predictors the 15 LEI scales and those dichotomous variables not being controlled.

4. Findings

Each of 32 null hypotheses, developed to test if scales from the LEI could predict SCA or SCAM scores, was rejected at the .05 level or less. In most cases, predictor variables accounted for 10% to 20% of the variance for self-concept. Favoritism appeared most frequently as a strong predictor, appearing in each of 16 attempts to predict SCA and 11 of 16 predictions of SCAM. Other variables which appeared with some frequency were: difficulty, cliquishness, and friction. Among the dichotomous variables included in the analyses the type of course was found to be a significant predictor of both SCA and SCAM. Results suggested that algebra students scored higher on both SCA and SCAM than non-algebra students. School size appeared as a predictor of SCA but not SCAM, with the weighting favored junior high school students. Sex appeared to be a very weak predictor, and having the teacher for a previous course was not a significant predictor at all.

5. Interpretations

This study provides evidence of the relationships between learning environments and self-concepts of academic ability. In addition, the results suggest that, for ninth-grade mathematics students, both general and mathematics self-concepts can be predicted from combinations of classroom environment characteristics. These results provide a rationale for further research to determine if it is possible to enhance self-concepts by providing suitable learning environments and thereby to increase academic achievement.

Critical Commentary

It is encouraging to find an interest in the influence of classroom environmental variables on student achievement and self-concept. Researchers have ignored for too long the influence of the learning environment. The author is to be commended for identifying variables to use as predictors of achievement which are different from such traditional variables as student and teacher backgrounds, student cognitive abilities, student attitudes, and teacher behavior. However, there are some serious limitations to this study which should be pointed out.

The recognition that environmental variables may play an important role in student achievement and self-concept carries with it a need to develop valid and reliable instruments to measure these variables. Although the author does not indicate, the LEI appears to require students to respond to questions about their perceptions of the classroom environment. Such responses may be unrealistic indicators of feelings with a very definite emotional bias. In the absence of information concerning the validity and reliability of the LEI, the reader should interpret the reported results with considerable caution. The same caution is given regarding the instrument used to assess the self-concept of academic ability. No description of the instrument is provided nor is validity and reliability evidence given.

Perhaps a more serious shortcoming of this report lies with the manner in which the sample and sampling procedures are described. The sample is described very loosely and the sampling method is not described at all. Also, no indication of how threats to internal and external validity were controlled is given.

The biggest problem in assessing this report rests not with the inadequate information about the instruments or the sampling procedures used, but with the manner in which the statistical analysis is reported. The author says that multiple regression was used but presents no summary data. Consequently, it is impossible to determine from this report whether or not the methods of analysis are valid and properly applied. Also, there is no way to know if the summary results which are reported accurately reflect the statistical analysis. Without further information about the data analysis, it is impossible to judge the merit of this study or to determine if the findings and interpretations are warranted.

Frank K. Lester, Jr.
Indiana University

EFFECTS OF A DEFINITION AND A VARYING NUMBER OF EXAMPLES AND NONEXAMPLES ON CONCEPT ATTAINMENT. Klausmeier, Herbert J.; Feldman, Katherine Voerwerk. Journal of Educational Psychology, v67 n2, pp174-178, April 1975.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by Thomas J. Cooney, University of Georgia.

1. Purpose

The study was designed to investigate the effects on concept attainment of four treatments: definition, a rational set of examples and nonexamples, a definition and a rational set, and a definition and three rational sets. A rational set was defined as a set of examples and nonexamples in which the examples vary widely in irrelevant attributes and the nonexamples differ from the examples in (if possible) only one relevant attribute at a time.

2. Rationale

There is evidence that sets of examples and nonexamples are most effective when they constitute a rational set. Further, studies have shown that the combination of a definition and a rational set is more effective in promoting concept acquisition than either a definition alone or a rational set alone. Most of this research, however, has dealt with college students and not with school-age children. Hence, the authors designed the present study to investigate the above question with school-age children.

3. Research Design and Procedure

"The subjects were 134 fourth-grade students from two elementary schools in a Wisconsin suburban community. Subjects were stratified into three groups (high, medium, and low) on the basis of grade equivalency scores obtained for reading on the Iowa Tests of Basic Skills administered during the fall of the school year. ...Experimental lessons and testing materials were administered in classroom groups."

The concept used was equilateral triangle. Four printed lessons in self-instructional booklets were designed as follows: (a) a definition and four pages of placebo information (Lesson D), (b) a rational set of three examples and five nonexamples--all drawings (Lesson RS_1), (c) a definition and the same rational set (Lesson DRS_1), (d) a definition and three different rational sets (Lesson DRS_{1-3}). Each booklet was five pages in length. Lesson DRS_{1-3} consisted of three 5-page booklets. Each group read three 5-page booklets. Those in treatments (a), (b), and (c), read two placebo booklets, while the control group read three placebo booklets. Total reading time was equated as much as possible by allowing subjects one minute per page. This time was judged to be sufficient.

The dependent measure was a classificatory task that required subjects to determine which instances in an array of 38 were examples and which were nonexamples. None of the instances had been presented in the booklets. They varied widely in difficulty. Booklets containing the dependent measure were distributed immediately upon completion of the lessons. Subjects were given as much time as necessary to complete the classificatory tasks.

"Conditions in which subjects read two placebo lessons and one treatment lesson (Conditions D, RS_1, and DRS_1) were counterbalanced so that half the subjects read the treatment lesson first and half read it third." Thus the experiment employed a 3x8 randomized block design with three levels of reading achievement (high, medium, and low) and eight experimental conditions.

4. Findings

A 3x3x2 analysis of variance (stratification, types of condition, order of presentation) was used to determine if the order of presentation (placebo vs. treatment) in lessons D, RS, and DRS, was significant. All effects due to order were judged nonsignificant and hence the data were collapsed across order of presentation.

This led to a 3x5 (stratification x conditions) analysis of variance. Only the main effect was significant. Differences in means for the stratifying variable were in the expected direction. The treatment effects were ordered as follows: $DRS_{1-3} > DRS_1 > RS_1 > D >$ control. Each treatment was significantly better than the control ($\alpha = .05$) and DRS_{1-3} group was significantly better than the D group. Differences among conditions D, RS_1, and DRS_1 were not significant. Achievement between the D and DRS_{1-3} groups increased from 76% to 88% correct ($p < .05$).

5. Interpretations

Subjects learned almost equally well from either a definition or a rational set, slightly more (although not significantly so) from both a definition and a rational set, and significantly more from a definition and three rational sets. Based on this and other studies, the authors believe the results would generalize to other concepts defined in terms of attributes. Further, since the results are consistent for fourth graders, sixth graders, and college students, the authors think it reasonable that the findings will transfer across grade levels. This has important implications for teaching students to discern examples from nonexamples of a concept. Further research was suggested for investigating the effects of presenting more than three rational sets.

Critical Commentary

This study provides useful information on how various teaching strategies affect concept attainment. In view of this study and others

cited by the authors, their claim that findings generalize across grade levels has merit. The Achilles heel of studies using self-instructional written materials, however, is the extent to which they have application in classroom situations.

The dependent measure is narrowly defined. One might argue that the superiority of Lesson DRS_{1-3} can be attributed to a practice effect. An interesting question is whether Lesson DRS_{1-3} is also superior in promoting other aspects of concept attainment. The authors' question concerning the effect on concept attainment of having more than three rational sets is also worthy of consideration.

Thomas J. Cooney
University of Georgia

SYNTAX, VOCABULARY, AND THE VERBAL ARITHMETIC PROBLEM. Linville, William J. School Science and Mathematics, v76 n2, pp152-157, February 1976.

[Note: This report is apparently based on Linville's doctoral dissertation, "The Effects of Syntax and Vocabulary upon the Difficulty of Verbal Arithmetic Problems with Fourth Grade Students" (Indiana University, 1969).]

Expanded Abstract and Analysis Prepared Especially for I.M.E. by Jeremy Kilpatrick, University of Georgia.

1. Purpose

The study was undertaken to determine whether the difficulty of verbal problems in arithmetic is significantly affected by variations in the syntax and vocabulary of the problem statement (with the numbers and operations held constant). Problem difficulty was also investigated with respect to the factors of pupils' sex, intelligence, and reading achievement.

2. Rationale

Changes in elementary school mathematics instruction have not produced pupils with the ability to apply their knowledge to the solution of problems. Although pupils may understand the computations, they have considerable difficulty solving verbal problems. Reading skills play a vital, if somewhat unclear, role in such problem solving.

3. Research Design and Procedure

One of four tests containing 10 word problems was randomly given to each of 408 fourth graders from 18 classes in 12 Indiana elementary schools. Five problems required addition and five required subtraction of two three-digit numbers. Problem statements were written at two "degrees" of syntax (easy syntax meant simple sentences only; difficult syntax meant complex sentences with one subordinate clause) and two levels of vocabulary (easy vocabulary meant at least 70 percent of the words were on the Thorndike-Lorge list of the first 1,000 words; difficult vocabulary meant at least 70 percent of the words were not on the list). Each test was composed of items with the same degree of syntax and level of vocabulary; the numbers and operation used in an item were constant across tests.

A panel of experts rated the test items for degree of syntax, clarity, and appropriateness of computation. A pilot study with 52 fourth graders yielded, among other information, KR-20 reliability coefficients of around .8 or above, except that the difficult-syntax/difficult-vocabulary test had a coefficient of .64.

The pupils' intelligence was measured by the SRA Primary Mental Abilities Test; their reading achievement by the reading comprehension subtest of the Iowa Tests of Basic Skills.

A 2x2x2x4 unweighted means analysis of variance was used, with sex, intelligence, reading achievement, and word problem test as factors. Scheffé's tests were used to explore the main effect for word problem test.

4. Findings

The mean scores (number correct) on the four word problem tests were significantly different ($p < .01$). The means on the two easy-vocabulary tests were significantly higher ($p < .01$) than the means on the two difficult-vocabulary tests. The means on the two easy-syntax tests were significantly higher ($p < .05$) than the means on the two difficult-syntax tests.

There was no significant difference between sexes in mean scores on the tests. Pupils with high intelligence had significantly higher ($p < .01$) mean scores than pupils with low intelligence, and pupils with high reading achievement had significantly higher ($p < .01$) mean scores than pupils with low reading achievement. There were no significant interactions among factors.

5. Interpretations

The investigator drew the following conclusions:

"1. Syntax and vocabulary level can both be determiners of difficulty in verbal arithmetic problems. Vocabulary level could be more crucial in determining success than syntax.

2. Boys and girls appear to do equally well in solving verbal arithmetic problems.

3. Pupils of higher ability can be expected to meet with considerably greater success in solving verbal arithmetic problems than pupils having less ability.

4. Pupils who have scored high in reading achievement can be expected to experience greater success in solving verbal arithmetic problems than pupils who scored low in reading achievement." (p. 157)

Critical Commentary

Sample items provided by the investigator show clearly that he varied not only the syntax and vocabulary but also the context, or setting, of the problem. Differences in problem difficulty cannot be attributed, therefore, to syntax and vocabulary alone.

Since the investigation was concerned with the pupil's ability to understand a verbal problem (apart from his or her ability to carry out the necessary computation), a more appropriate dependent variable would have been the number of problems in which a correct computation was set up.

No mean scores or standard deviations are reported. The use of an unweighted means ANOVA suggests that the cell sizes were unequal, but these numbers are not given either. A more efficient analysis would have used syntax and vocabulary as separate factors. Also, the pupil's class or school ought to have been considered in the analysis.

In addition to having no explicit link to theory or previous research, the study was poorly designed and reported. It adds little to our knowledge.

Jeremy Kilpatrick
University of Georgia

FORWARD TRANSFER OF DIFFERENT READING STRATEGIES EVOKED BY TESTLIKE EVENTS IN MATHEMATICS TEXT. Mayer, Richard E. Journal of Educational Psychology, v67 n2, pp165-169, April 1975.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by Clyde A. Wiles, Indiana University Northwest.

1. Purpose

The purpose of this study was to investigate forward transfer effects of testlike events (written questions and behavioral objectives) on what is learned from written materials involving probability concepts.

2. Rationale

Previous studies, none of which involved mathematical content, indicated that students who read passages differing only in the types of questions inserted in the text tended to perform on post-tests in predictable ways. This effect has been accounted for by both forward and backward models. The forward model explains differences in performance on the basis of learning sets established by the preceding questions. The forward transfer model is contrasted with a backward effects model which postulates that questions serve primarily to review or rehearse already learned material, or as a practice for answering questions of specific type. The study at hand was designed to investigate the forward transfer model further and to do so within the context of mathematical concepts and problem solving.

3. Research Design and Procedure

The forward transfer effects of ten questioning strategies were investigated. These involved five levels of "question type" (defining, calculating, model, all, and none) and two levels of "question conditions" (preview, as behavioral objectives with no required response; and review, as usual questions that were to be answered). While all ten strategies are called questioning strategies, half of them involve behavioral objectives stated in a preview condition rather than questions.

The subjects were 80 college sophomores who participated in the study as a course requirement. After insuring that all subjects were (a) naive with respect to the content and (b) able to compute, they were separated into 10 groups on the basis of SAT-M scores so that the average scores for all groups were equivalent.

All treatments contained the same eight printed instructional passages. The first six passages were either preceded by a sheet containing objectives of the specified type or were followed by a sheet containing questions of the specified type. The "none" groups received neither questions nor objectives for the first six passages. The pattern for the last two passages was the same for all groups. Each of these passages was followed by three sets of four questions; one set was characterized as defining, another as calculating, and the last as modeling (application). The scores for these sets were the criterion variables. The time required by each subject to

read each passage was also recorded. Materials were sequenced and collected in such a way that questions or objectives were not in the possession of subjects at the same time as the passages were.

4. Findings

The main effect of question type was significant in the multivariate sense. The two groups receiving all types scored the highest and the groups receiving none scored the lowest. The interaction between question type in treatments and question type on the post-test was also significant, with differences in the expected directions. While there were significant differences in overall performance from passage 7 questions to passage 8 questions, there were no significant interactions with treatments. The main effect of question conditions (preview versus review) was not found to be significant nor were any interactions involving conditions significant.

While no significance tests of differences in reading times were reported, the average reading times for treatments ranged from 164 seconds for the All group in the review condition to 203 seconds for the Definition group also in the review condition.

In summary:

(a) Different questions or objective types inserted into the learning experiences of the first six passages resulted in predictable differences in what was learned from passages 7 and 8.

(b) Behavioral objectives presented as a preview to reading and which required no overt response other than reading had the same forward transfer effect as questions stated as a review and which asked for written responses.

5. Interpretations

The observed differences in performance were accounted for in terms of forward transfer. Since the last two passages involved content identified as independent of any specific information given in the earlier passages, the effect cannot be attributed to the backward influence of prior questions. While backward processes of questions are not ruled out, forward transfer of questions or objectives is thus shown to be both significant and a potentially important factor in learning.

The fact that the preview condition groups did not practice answering any questions, but were yet indistinguishable from the review condition groups who did, was taken to mean that the effects could not be attributed to practice in answering certain question types.

In sum, it was felt that different groups learned different things from the same passages because they adopted different reading strategies, evidently on the basis of the preceding 6-passage experience. Differences in average reading times, while not tested for significance, were shown to be consistent with this interpretation.

Post hoc analysis indicated that Definition group subjects not only excelled on definition items on the post-tests, but they also performed quite

well on calculation items. The Calculating groups in turn exhibited a similar pattern. However, both Definition and Calculating groups performed poorly on model (application) items. In contrast, the Model groups not only excelled on the model items but did very well with the Calculating and Definition items as well.

The author implies, therefore, that "testlike events that emphasize understanding do not limit hypotheses concerning the goal of learning as much as an emphasis on rules or computation." This implication was noted as being consistent with earlier studies.

Critical Commentary

This study is a model of careful design. While one may criticize the arbitrary assigning of subjects to treatments on the basis of SAT-M scores rather than randomly, this did control a strong confounding variable and still left a reasonable claim for group equivalence. It is not a criticism of the study, but it should be noted that mathematics education was not the focus of the study. Mathematics provided an unambiguous context for the study and contributed to the generalizability of the findings. Related studies used non-mathematical contexts.

The study is seen to have the following implications for mathematics education:

a. The role of questions in affecting what a student is likely to learn from succeeding experience is established. The nature of "quizzes" and "warm-up exercises" used to introduce lessons and units needs to be carefully considered.

b. The ability of well-stated behavioral objectives to direct the learning of students as effectively as questions is important. This suggests that careful attention should be given to selecting objectives that do identify what is most important, and that communicating these to students does indeed make a difference.

c. The apparent equivalence of focusing upon memory tasks or upon computational tasks and the resulting failure of the learner to attend to understanding (applications in this case), coupled with the observation that application tasks may produce learning of all three tasks, is quite important. The nature of the analysis and the specialness of the population make this more of an implication for further research than for practice at this time, however.

Other questions for research are indicated by this study. Note first that the subjects were college sophomores at a major university, the content was abstract/symbolic, and the instructional context was entirely printed.

a. Would the same forward transfer of testlike events be evident in younger, less symbolically sophisticated learners?

b. While the effects were induced by written materials which the learners read, would other mediums serve as well?

c. The experience of six preceding passages was shown to be sufficient to induce differences in learning from the same material; could this be reduced without weakening the effect?

d. Suppose a learner set had been induced through some sufficient experience; how long would this set persist? After having once established a learner set, might it be possible to reestablish it by a rather simple stimulus?

Related to the establishing or triggering of a learner set, the study at hand involved sophisticated learners who would be reasonably expected to have a rather large repertoire of learning strategies. One wonders how this whole process may be related to less sophisticated and/or less capable learners.

Clyde A. Wiles
Indiana University Northwest

IS CONSERVATION OF NUMBER A NECESSARY CONDITION FOR MATHEMATICAL UNDERSTANDING? Mpiangu, Benayame Dinzau; Gentile, J. Ronald. Journal for Research in Mathematics Education, v6 n3, pp179-192, May 1975.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by James J. Hirstein, University of Illinois.

1. Purpose

The hypothesis that conservation of number is a necessary condition for acquisition of numerical concepts is considered basic to Piaget's theory. This study was intended to test that hypothesis.

2. Rationale

Previous conclusions regarding the hypothesis in question are seen to infer causal relationships from high correlations between conservation and arithmetic achievement. Furthermore, the authors take issue with the usual practice of treating conservation as a dichotomous variable. At least two benefits are gained by treating conservation as a continuous variable: regression techniques are made available and the traditional "intermediates" on conservation (that is, students not classified as conservers or non-conservers) are not excluded from the study.

3. Research Design and Procedure

The study used a pretest-treatment-posttest design. An arithmetic pretest of 15 items was administered individually to 137 kindergarten students in two schools. Twenty-one subjects were discarded, leaving 116 subjects in the study. These remaining subjects were given an eight-item pretest on conservation of number, and assigned to experimental and control treatment groups of about six subjects each.

The experimental treatment group was given arithmetic training on the following topics: (a) counting forward and backward; (b) finding the missing number in a sequence; (c) determining numbers which come before, just before, after, just after, and between other numbers; and (d) answering and justifying answers for simple addition and subtraction items. The control treatment group played a card game involving the identification and discrimination of animal pictures. Each group met for ten sessions of 15-20 minutes.

Following the treatment sessions, parallel forms of the conservation and arithmetic tests were administered to all children. The arithmetic posttest also contained two items designed to measure more advanced material than that included in the experimental treatment.

4. Findings

Conservation Posttest and Arithmetic Posttest were the measures of the dependent variables of the study. Multivariate analyses of covariance

were carried out using Arithmetic Pretest, Conservation Pretest, and Both Pretests as covariates. Each analysis showed Treatment to be a significant effect ($p < .0001$) in favor of the experimental group. Univariate analyses of covariance for each dependent variable showed the experimental treatment group to be far superior to the control group on the Arithmetic Posttest and only moderately superior to the control group on the Conservation Posttest.

A comparison of treatment groups was conducted using the regression of posttest scores on the Conservation Pretest score. Both the Arithmetic Posttest scores and the Conservation Posttest scores showed a homogeneity of regression (parallel regression lines) for the two treatment groups.

A second examination of the data was conducted treating conservation as a dichotomous variable. Those subjects scoring the minimum on the Conservation Pretest were considered non-conservers and those scoring the maximum were considered conservers. Intermediate subjects were discarded. Using Arithmetic Gain (posttest minus pretest) as the dependent variable, the analysis showed a significant Treatment effect, but no significant Conservation effect and no significant Treatment-Conservation interaction.

5. Interpretations

If conservation of number were a necessary condition for the learning of arithmetic, then conservers would benefit from arithmetic instruction while non-conservers would not. Homogeneity of regression of Arithmetic Posttest on Conservation Pretest for experimental and control groups implies that non-conservers benefited essentially the same from instruction as conservers. The high significance of Arithmetic Posttest compared to the modest significance of Conservation Posttest indicates that "...the treatment effect acted mainly on arithmetic performance without substantially altering conservation performance" (p. 188). These results are interpreted as evidence that conservation is not a necessary prerequisite for mathematical understanding. The authors infer that, in the derivation of arithmetic teaching strategies, "...conservation of number should not be regarded as a crucial concept" (p. 191).

Critical Commentary

Gains in arithmetic scores were quite dramatic for ten 20-minute sessions with kindergarten children. A note about the time lag of posttests from treatment would have been helpful. The conclusions would be more credible if the results were known to hold for a relatively long retention period.

The treatment of conservation as a continuous variable raises some difficulties. Although development itself is a continuous process, developmental phenomena are often traced using characteristics which are "present" or "absent." The large number of "intermediates" in conservation studies supports the authors' contention that conservation is not an "all-or nothing" phenomenon. But even if the acquisition of conservation is a continuous process, it is doubtful that a score on a set of classical conservation items captures the continuity of the process. Typically, these tests indicate a tendency (or lack of tendency) to respond to

configurational or numerical aspects of a transformation. Consequently, past and future studies should not be faulted for treating conservation as a discrete variable.

The report of this study shows no effort to ascertain the internal mechanisms which a child uses in the acquisition of number concepts. In this omission the authors deny one of the basic tenets of those who would derive arithmetic instructional strategies from Piagetian theory, that the selection of experiences for learning mathematics be made on the basis of the method of attack available to the child. Knowing what topics can be learned by the pre-operational child is of great importance to developers of early childhood curricula. Of equal importance is the knowledge of the effects of various instructional strategies on children whose cognitive structures are susceptible to change. When different approaches are found to have different effects on children of varying cognitive capacities, the concepts used to identify an appropriate strategy should be regarded as crucial.

James Hirstein
University of Illinois

THE ROLE OF STRATEGIES FOR TEACHING PUPILS TO SOLVE VERBAL PROBLEMS.
Richardson, Lloyd I. Arithmetic Teacher, v22 n5, pp414-421, May 1975.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by Aaron D. Buchanan, SWRL Educational Research and Development, Los Alamitos, California.

1. Purpose

The primary purpose of the research was to show that teaching pupils specific, step-by-step strategies for solving verbal problems will improve their performance. The purpose of the article itself was to provide classroom teachers with alternative strategies to use during instruction.

2. Rationale

Verbal problem solving is a complex behavior where pupils must learn to process the information given in a problem by (1) identifying the substantive statements, (2) organizing them so they converge on "what must be done" to solve the problem, and (3) translating "what must be done" into some sort of mathematical relationship that expresses the solution. Pupils need generalizable strategies for processing the information in verbal problems, because of the complexities of language and prose. Often the same underlying mathematical relationship applies to problem statements that, on the surface, appear to be quite different in structure.

All too frequently, strategies for solving verbal problems are not taught to pupils because teachers are not aware that specific strategies exist. Presumably, if teachers become better acquainted with such strategies so they can include them as part of regular instruction, better student performances on verbal problems will be the result.

3. Research Design and Procedure

Two studies are reported. One study was conducted by the author (Richardson, 1973), and another by the author and two colleagues (Bassler, Beers, and Richardson, 1972). Both studies involve two strategies for problem solving. In the Step Method, the procedure was to: (1) read the problem, (2) decide what question is being asked and represent the unknown with a variable $\square$, (3) consider other information in the problem and how it relates to the unknown, (4) write a number sentence expressing the given relationship, (5) find the number that makes the sentence true, and (6) check the answer. In the Translation Method, the procedure was to translate each phrase in the problem, line by line, into mathematical "symbol" statements.

For example:

$H + C = 64$

A TRAY OF HAM AND CHEESE SANDWICHES CONTAINS 64 SANDWICHES IN ALL.

$33 = C$ $\qquad$ $H = ?$

THERE ARE 33 CHEESE SANDWICHES. HOW MANY HAM SANDWICHES ARE ON THE TRAY?

The key element in each phrase is represented by a variable, preferably the first letter in a key word. The variable representing an unknown becomes a question mark (?) or a □ (e.g., H = □). The resulting set of symbol statements is then translated into a number sentence containing the unknown (□ + 33 = 64), and a number is found which makes the sentence true (□ = 31). Finally, the result is checked (33 + 31 = 64).*

The first study (Richardson, 1973) involved 12 preservice teachers and 247 ninth- and tenth-grade pupils. First, the teachers were randomly assigned to two treatment groups. In one group, teachers were given four hours of instruction dealing with the two methods for solving verbal problems. In the other group, which served as a control for previous experience, teachers were given four hours of discussion on enrichment topics that specifically excluded verbal problems. Each teacher was then given a unit to teach on verbal problem solving which involved a pretest, three days of instruction, and a posttest developed by the author. The 247 pupils in the study made up twelve regular classes in general mathematics.

In the second study (Bassler, Beers, and Richardson, 1972), a large number (N = ?) of ninth-grade pupils was given seven days of instruction on either the Step Method or the Translation Method, followed by a posttest and then a retention test four weeks later.

4. Findings

The primary data in the first study were pupils' scores on the posttest, where the single criterion was whether pupils had a correct solution to test items. An analysis of covariance was conducted with an adjustment for scores on the pretest. For the analysis the author randomly chose data from 15 pupils in each class who had successfully completed the unit. (These pupils must have taken the pretest, posttest, and been present for at least two days of instruction.) The results showed that mean posttest scores for pupils in the experimental group (whose teachers had pretraining on the two methods for solving verbal problems) were significantly higher than for pupils in the control group, but only at the .10 level. Differences in pupils' scores based on other factors, such as whether their teachers had high or low grade point averages, were not significant.

In the second study, data from 96 pupils were analyzed according to: (1) whether they wrote correct equations in their procedures, and (2) whether they had the correct solution to the overall problem. There were no significant differences for either problem-solving method on the posttest. This result was true for correct equations as well as for final solutions. The retention test showed no significant differences between problem-solving strategies, but solution scores on the retention test were significantly higher ($p < ?$) than solution scores on the posttest.

---

*This example is a simplified version of an example given in the article.

5. Interpretations

It was generally concluded that instruction on a specific strategy was better than a more general approach to solving verbal problems, but, as far as the two strategies involved in these studies were concerned, it did not make much difference which one was used. The author suggested a combination of the Step and Translation Methods as one that teachers might want to try, particularly in junior high school classes.

Critical Commentary

The introductory material in this report goes into commendable detail in describing to teachers some of the complexities involved in verbal-problem solving. For teachers who are having difficulty teaching verbal problems, and many are, the report offers a number of good suggestions for focusing instruction on specific learning problems.

As a report on research, this article leaves more to be desired. In the first study, the treatments are described carefully enough, as they should be in a journal aimed primarily at classroom practice, but it is hard to tell how meaningful the results are. Two additional details would have helped immensely. First, there is no way to tell what strategies were used by teachers who were not given pretraining on the two methods for solving verbal problems. Since pupil results based on whether or not teachers were pretrained were only marginally significant ($p < .10$), one can only wonder if strategies used by the two groups of teachers turned out to be all that different. Second, without knowing how many items were on the posttest, there is no way to tell whether the results have any practical significance. The mean score for the experimental group was 14.39, while the mean for the control group was 12.98. If the post-test had 15 items, then both scores were "good," whether they differed significantly or not. If the posttest had 30 items, then the results are a "plague" on the houses of both treatments.

The report on the second study really had little to say to either teachers or researchers. So few technical details were provided about the purpose, sample size, nature of the treatments, and so on, that many questions are raised, and few, if any, are answered.

It would be easy to fault the author for omitting so many details that are important for readers interested in the design of research, but it would not be fair. The Arithmetic Teacher is not a journal intended for researchers. Since February 1971, the editors have specifically requested articles that provide interpretations and applications of research which can be used directly in the classroom. Whether this factor accounts for so many missing details in this article is hard to say. The suggestions given by the author for improving verbal-problem solving are good ones, but it is not at all clear that they are supported very strongly by the research they accompany.

REFERENCES TO ORIGINAL STUDIES

Bassler, Otto C.; Beers, Morris; Richardson, Lloyd. Comparison of Two Instructional Strategies for Teaching the Solutions of

Verbal Problems. United States Department of Health, Education and Welfare, Office of Education, National Center for Educational Research and Development (OEG-4-72-002), December 1972. (ERIC: ED 075 260).

Richardson, Lloyd I. Student Achievement in Solving Verbal Problems as Related to Teacher Preparation. Ph.D. dissertation, George Peabody College for Teachers, 1973.

Aaron D. Buchanan
SWRL Educational Research and Development
Los Alamitos, California

CRITERIA FOR ASSESSING THE EFFECTIVENESS OF TEACHERS OF SECONDARY SCHOOL MATHEMATICS. Robitaille, David F. Journal for Research in Mathematics Education, v6 n2, pp77-87, March 1975.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by Arthur F. Coxford, University of Michigan.

1. Purpose

To classify teachers as highly or minimally effective. Then to compare the classified groups of teachers with respect to selected background variables and test scores and to compare the frequencies of selected classroom teaching behaviors for the two groups.

2. Rationale

The author suggests that little firm evidence exists regarding the relationship between effectiveness in teaching and a variety of other variables. In fact no agreed-to definition of "effectiveness" is found in the literature or in this report. Some results were cited indicating that students of junior high school mathematics learned more with "indirect teachers" than with "direct teachers". Since indirect teachers were those who solicited ideas, expanded student-offered ideas, and encouraged and supported students, the present study sought to develop an instrument which would measure the extent of pupil participation in the lesson. The contention was that highly effective teachers would promote greater pupil participation than minimally effective teachers.

3. Research Design and Procedures

A set of 167 secondary (grades 8-12) mathematics teachers were independently rated by four judges. Each judge rated each teacher as to effectiveness (with no agreed-to definition of the word). The ratings were added, thus identifying 23 highly effective and 20 minimally effective teachers.

Each of the 43 teachers was assessed on "contemporary mathematics", "teaching situation reaction", "teacher attitude", "adjective checklist", and "teacher opionionaire". In addition five other background variables were defined: age, years of teaching, mathematics background, percentage of assignment devoted to teaching mathematics, and years since a degree was granted. These data were used in the exploratory phase in which it was determined which variables best predicted group membership, using discriminant analysis.

The author developed the Pupil Involvement Check List which was designed to measure the degree to which teachers involved students in the lesson. This instrument was used by the author in 50 randomly made visits, 25 to each type of classroom teacher. These data were used in the classroom observation phase in which scores of the two groups were tested by student's *t*.

4. Findings

For the exploratory phase, the first variable entered (contemporary mathematics) correctly placed 17 of 23 highly effective teachers and 16 of 20 minimally effective teachers. By the time eleven variables were entered these figures increased to 22 of 23 and 19 of 20. In further discriminant analyses, all but 3 highly effective teachers had conditional probabilities of over 0.9 of being classified as highly effective; similarly all but 2 of the minimally effective teachers had similarly high conditional probabilities of being classified as minimally effective on the basis of the background data collected.

The results of the classroom observations showed that the highly effective teachers scored significantly higher on the total Pupil Involvement Check List and on the following three items: "asks 'thought provoking' questions", "encourages a pupil to contribute to the lesson", and "praises or encourages a pupil or the class."

5. Interpretations

The author concludes that "...there are measurable and observable differences between highly effective and minimally effective teachers of secondary school mathematics". Of particular importance seems to be the concept of teacher encouragement of student participation in a lesson. Furthermore, the four-judge rating procedure "...proved to be at least as effective a means of classifying teachers into dichotomous groups as were the..." use of the background data collected in the exploratory phase. Finally, "...the prevalence of such behavior (teacher encouragement) was found to be a significant indicator of teacher effectiveness."

Critical Commentary

The educational significance of this study is not at all clear. Neither is its theoretical value. A major weakness is the lack of a definition of "effective teacher". Even though the four judges showed remarkable agreement in their ratings, there is no assurance that the teacher they perceive as effective actually is. This reviewer would suggest that "effective teacher behaviors" differ for different mathematical topics and for students of different capabilities. Any attempt to study effectiveness that does not take such considerations into account is doomed from its inception.

A second criticism of the work is the lack of a clear definition of the problem. There are no clearly stated hypotheses nor questions to answer. Rather copious data were collected and analysed, seemingly with little clear purpose. The author suggested that "teacher encouragement" was an indicator of teacher effectiveness, yet the only conclusion which could be drawn from the data was that "effective teachers" were rated significantly greater on this aspect than were "ineffective teachers". The lack of clear purposes led to this error and the general meandering nature of the work.

As difficult as it may be, the effectiveness of a teacher must be related to learning and to the attitudes of students. Only when this is done will it be profitable to seek factors characteristic of an "effective teacher".

Arthur F. Coxford
University of Michigan

THE EFFECTIVENESS OF FOUR READINESS TESTS AS PREDICTORS OF FIRST GRADE ACHIEVEMENT. Telgedy, Gabriel A. Psychology in the Schools, v12 n1, pp4-11, January 1975.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by Merlyn J. Behr, Northern Illinois University.

1. Purpose

The study was designed to examine the effectiveness of selected readiness tests to predict success in first grade.

2. Rationale

The study was conducted in the context of the question about whether the progress of young children in their first year of formal education is dependent upon their "readiness skills." Although readiness tests have been developed, their predictive validity is frequently low. Moreover, independent of the fact that the construct content validity of the tests appears adequate, factor analysis often reveals overlaps among subtests; this results in tests longer than necessary. The question of whether readiness tests are better predictors of first-grade success than teacher ratings was cited. Also the question of whether or not a battery of readiness tests provides better predictive power than a single readiness test was of interest.

3. Research Design and Procedure

Four readiness tests [Screening Test of Academic Readiness (STAR), Bender Gestalt TEst (BGT), First Grade Screening Test (FGST), and Metropolitan Readiness Test (MRT)] were administered to 56 children (28 boys and 28 girls) at the end of their kindergarten year. One year later, the same children were administered the achievement tests [Wide Range Achievement Test (WRAT) and Gray Oral Reading Test (GORT)]. The Ss consisted of two kindergarten classes of two schools representative of the Windsor Separate School system. The children of the two schools combined provided a sample of advantaged and disadvantaged pupils from upper, middle, and lower socioeconomic levels.

4. Findings

Based on computed correlation coefficients, it was concluded that the MRT was the best predictor of all five criterion scores--WRAT Reading, Spelling, and Arithmetic and GORT Passage and Comprehension. The STAR and MRT were very close in predictive value and clearly superior to the BGT, FGST and teacher ratings. The teacher ratings were as good, or better than, the BGT and FGST in predicting Reading, Spelling and GORT Passage and Comprehension scores. The Letters and Number subtests of the STAR were as predictive as the entire test composed of eight subtests.

The interested reader will find reported in the paper results of t-tests which were presented to determine whether differences in predictive value indicated above attained a level of statistical signficance.

Results of regression analysis suggested that combining all five predictors into a larger battery would not significantly improve predictive validity. Results of stepwise multiple regression for maximizing R^2 improvement, indicate that: the best set of four predictors for WRAT Reading were STAR Letters, GHT, HFO, and MRT Alphabet; for WRAT Spelling, STAR Lectures, BGT, FGST, and MRT Alphabet; for WRAT Arithmetic, FGST Picture Completion and Copying, and Teacher Prediction of Arithmetic; for GORT Passage, STAR Letters, GBT, HFO, MRT Alphabet; and for GORT Comprehension, BGT, MRT Alphabet, Teacher Prediction of Reading and Spelling.

5. Interpretations

The authors observe that different achievement criteria could result in higher predictive validity for the different readiness tests. The authors' findings confirm earlier results that teacher ratings are quite effective in predicting first-grade success. The advantage of standardized tests, they claim, lies in their ability to point very clearly and impartially to the specific deficits for which remedial work should be initiated. While teacher ratings include consideration of a child's social and emotional status, these frequently fail to specify the particular skill that may hinder a child's future progress. Moreover, stepwise regression results have shown teacher ratings to possess significant predictive value and, thus, a recommendation is made to use teacher ratings in conjunction with standardized tests. The authors cite other research concerning training on readiness skills found to be correlated with first-grade achievement which indicate that such training results in higher readiness test scores, but not in improved post-achievement test scores.

The results of this research have several implications for professionals in education. They suggest that STAR or MRT, coupled with teacher observations and the BGT, can provide a school system with an adequate screening readiness battery.

Critical Commentary

The study gives some information about the usefulness of available readiness tests for predicting first-grade achievement. The authors are to be commended for including teacher ratings as a potential indicator of first-grade achievement. On the other hand, the question of why they did not use teacher ratings of first-grade achievement occurs. In view of the fact that standardized tests were used both for readiness and achievement measures, the study may simply give data to suggest that standardized tests administered at the end of kindergarten are correlated with standardized tests given at the end of first grade. One might conjecture that had teacher ratings been used as a criterion for first-grade achievement, then the teacher readiness ratings would have had the best predictive value. This concern seems to be particularly relevant in view of the fact that its publication occurs in the midst of a major concern about the appropriateness of standardized testing.

This study is not exempt from a common concern about generalizability of results. Two intact classes, two teachers, and 58 students is a small sample from which to generalize results for a correlational study. There is also concern about the statistical tests used in the study. One set of data were subjected to a number of analyses: correlation analysis, multiple t-tests, regression analysis, and stepwise regression analysis.

Absent from the study is any indication about whether these readiness scores are correlated to individual differences such as sex, race, socio-economic level, or intelligence. No information is available about whether children with different readiness profiles achieve differently in various classroom environments.

Merlyn J. Behr
Northern Illinois University

THE EFFECT OF MATHEMATICS INSTRUCTION USING MANIPULATIVE MODELS ON ATTITUDE AND ACHIEVEMENT OF PROSPECTIVE TEACHERS. Warkentin, Gary. Journal for Research in Mathematics Education, v6 n2, pp88-94, March 1975.

Expanded Abstract and Analysis Prepared Especially for I.M.E. by Larry Sowder, Northern Illinois University.

1. Purpose

The primary purpose of the study was to compare the effects of (1) a laboratory-and-models approach and (2) a lecture-and-textbook approach on the achievement and attitude of prospective elementary school teachers in a required college mathematics content course. Also reported were how attitudes changed under the two approaches and how scores on a standardized arithmetic test were correlated to final grades in the course.

2. Rationale

Several surveys have revealed that prospective elementary teachers, on the average, are not strong in mathematics. Nor does the usual coursework in college mathematics seem to improve their attitudes appreciably. Some theorists feel that work with symbols should be preceded by work with concrete materials or "real" experiences. Perhaps a course using manipulative materials in a laboratory setting can improve attitudes and lead to greater achievement than the usual course.

3. Research Design and Procedure

Subjects were students in a theory of arithmetic course, with 6 sections (N = 149) undergoing the experimental laboratory-and-models approach and 9 control sections (N = 197) using the customary lecture-and-textbook approach. Two pretests were given: (a) a 20-item Likert-type attitude scale (Dimensions: liking/disliking mathematics, tendency to avoid/engage in mathematical activity, belief that one is good/bad in mathematics, belief that mathematics is useful/useless), and (b) the arithmetic section of the advanced level of the Stanford Achievement Test. The groups were comparable on these measures.

The treatment period was a full semester, with both groups considering the same subject matter. Students in the experimental sections worked in groups of four with manipulative materials (colored rods, geoboards, numeration system blocks, balances) and with task cards and activity sheets based on the materials.

The two posttests were (a) the attitude questionnaire used as a pretest and (b) a 50-item achievement test based on the textbook used by the control group.

4. Findings

a. On the attitude measure, a t-test showed that the mean for the experimental group (66.3) was significantly ($p < .01$) different from the mean for the control group (62.6). For the experimental group, the posttest mean of 66.3 was significantly different from the group's 61.6 pretest mean ($p < .01$). Pre- to posttest means did not change significantly for the control group (61.5 to 62.6).

b. On the achievement measure, the control group (33.4) was significantly ($p < .01$) different from the experimental group mean (30.6).

c. The correlations between the Stanford Achievement Test arithmetic score for the final grade were .64 for the experimental group and .58 for the control group, both coefficients being significantly different from 0.

5. Interpretations

a. The laboratory-and-models approach seemed to have a positive effect on the students' attitudes toward mathematics.

b. The control sections were able to cover more material than the experimental sections. This fact may explain the control group's apparent superiority on the achievement test.

c. Further research should investigate the students' retention, their success in mathematics methods courses, their success in the classroom, and the optimal amount of time that the students should spend with manipulative materials.

Critical Commentary

a. Even if this study had given clear support to the experimental treatment, one could still be in a quandary as to what deserved the credit. Such credit might have been due to the activity-oriented work or to the hands-on use of manipulative materials, as is implied by the phrase "laboratory-and-models approach." But perhaps only the use of small groups deserves the credit (this would nicely explain the attitude results). Or perhaps the experimental written materials were just not as good as the control group's textbook (which might explain the achievement results). Perhaps the control group teachers were better content-explainers (no information was given on who the teachers were; this variable may have been controlled), while the greater teacher-student and student-student interaction in the laboratory setting led to the change in attitude (the role of the teachers in the experimental classes was not explained). In sum, it is not clear what led to any observed differences.

b. Quibbles: Presumably there was some certainty that students were assigned to treatments randomly. Also, there was no information on mortality (attrition). Could some other design have tested for the possible sensitizing effect of the attitude pretest?

c. What was the nature of the tests taken during the semester by the two groups? Mere test-format familiarity could have given the control group an advantage on the achievement test.

d. Since the experimental group did not cover all the material on the achievement test, it would have been informative to have compared performance on only the material covered by both groups. What was the reliability of the the achievement test? Could one identify sub-tests on the achievement test (e.g., concepts, applications,...)? If so, how did performances of the two groups compare on them? For example, in some unpublished work currently being replicated, Barnett and Eastman have tentative indications that hands-on work with manipulatives may enhance concept attainment more than work with pictures of manipulatives does (for college students in methods classes).

e. Might the attitude differences have all arisen from one dimension? Is there any information on the reliability of the attitude scales?

Larry Sowder
Northern Illinois University

MATHEMATICS EDUCATION RESEARCH STUDIES REPORTED IN RESEARCH IN EDUCATION

April - June 1976

ED 115 020 Toniolo, Thomas A.; Hooper, Frank H. Micro-Analysis of Logical Reasoning Relationships. Conservation and Transitivity. Technical Report No. 326. 74p. MF and HC available from EDRS.

ED 115 040 Thurlow, Martha L.; And Others. Measurement of Length Unit: A Formative Evaluation. Research Report No. 85. 111p. MF and HC available from EDRS.

ED 115 448 Boulanger, Ferdinand David. The Effects of Instruction in the Concept of Speed and Proportions on Children in the Third Grade. 155p. Available from University Microfilms (74-2243).

ED 115 480 Trent, John H. Comparative Needs of Elementary and Secondary In-Service Teachers and College Pre-Service Students for Metric Education. 25p. MF and HC available from EDRS.

ED 115 486 Resnick, Lauren B. Task Analysis in Instructional Design: Some Cases from Mathematics. 51p. MF and HC available from EDRS.

ED 115 490 Romberg, Thomas A. (Ed.) Four Studies on Aspects of Assessing Computational Performance. Technical Report No. 297. 115p. MF and HC available from EDRS.

ED 115 491 Romberg, Thomas A.; Glove, Richard. Process Models for Predicting the Difficulty of Multiplication Problems Using Flow Charts. (Part 1 of 2 Parts). Technical Report No. 337. 82p. MF and HC available from EDRS.

ED 115 492 Romberg, Thomas A.; Glove, Richard. Process Models for Predicting the Difficulty of Multiplication Problems Using Flow Charts. (Part 2 of 2 Parts). Technical Report No. 337. 47p. MF and HC available from EDRS.

ED 115 493 District of Columbia Public Schools, Washington D.C., Dept. of Research and Evaluation. Learning Through Selected Symbols, ESEA Title III Evaluation, Final Report. 21p. Available (on loan) from ERIC/SMEAC, The Ohio State University, 1200 Chambers Road, 3rd Floor, Columbus, Ohio 43212.

ED 116 907 Drew, William Donald. Development of a Set of Teaching-Learning Guides Unifying Selected Areas of Secondary School Mathematics and Physics. 303p. Available from University Microfilms (74-10,046).

ED 116 911 Fournier, Raymond H. Development and Implementation of Computerized Monitoring System in Mathematics Grades 4, 5, 6. 144p. MF and HC available from EDRS.

ED 116 912 Carlson, Gaylen R. Specific Concepts of Space and Time in Children from Grades One Through Six. 18p. Available (on loan) from ERIC/SMEAC, The Ohio State University, 1200 Chambers Road, 3rd Floor, Columbus, Ohio 43212.

ED 116 917 Shapiro, Bernard. The Notebook Problem. Report on Observation of Problem Solving Activity in USMES and Control Classrooms. 26p. MF and HC available from EDRS.

ED 116 918 Shapiro, Bernard; Aiello, Thomas. USMES Evaluation Report on Classroom Structure and Interaction Patterns, 1972-73 USMES Evaluation Program. 19p. MF and HC available from EDRS.

ED 116 926 Robie, Douglas E.; And Others. Calcumate Part I: Effects of Using a Mechanical Device to Teach Subtraction with Borrowing to School Children (Abbreviated Manuscript). 17p. MF and HC available from EDRS.

ED 116 938 Begle, E. G. Ability Grouping for Mathematics Instruction. A Review of the Empirical Literature. SMSG Working Paper No. 17. 55p. MF and HC available from EDRS.

ED 116 940 Poll, Dwayne C.; Allegra, Michael. The Individualization of Math in Grades Five and Six: An Experimental Program. 19p. MF and HC available from EDRS.

ED 116 966 Callahan, Leroy G.; Glennon, Vincent J. Elementary School Mathematics: A Guide to Current Research, Fourth Edition. 2[illegible]p. MF available from EDRS. HC not available from EDRS.

ED 116 967 McMaster, Mary Jane. Differential Performance of Fourth- Through Sixth-Grade Students in Solving Open Multiplication and Division Sentences. (Part 1 of 2 Parts). Technical Report No. 375. 127p. MF and HC available from EDRS.

ED 116 968 McMaster, Mary Jane. Differential Performance of Fourth- Through Sixth-Grade Students in Solving Open Multiplication and Division Sentences. (Part 2 of 2 Parts). Technical Report No. 375. 77p. MF and HC available from EDRS.

ED 117 142 Higgins, Paul S. The Mathematics Component of the Minneapolis Schools' 1973-74 Emergency School Aid Act Project. An Evaluation. 48p. MF and HC available from EDRS.

ED 117 499 Small, Charles; And Others. [Arizona] Field Test Report. Vol. 12. Construction Industry Related Math. 1974-75. 48p. MF and HC available from EDRS.

ED 117 861 Powell, Marjorie. Teacher Competencies: California Beginning Teacher Evaluation Study. 19p. MF and HC available from EDRS.

ED 118 359 Hadar, Nitsa Boneh. Children's Conditional Reasoning: An Investigation of Fifth Graders' Ability to Learn to Distinguish Between Valid and Fallacious Inferences. 420p. MF and HC available from EDRS.

ED 118 417 Gonchar, Arthur J. A Study in the Nature and Development of the Natural Number Concept: Initial and Supplementary Analyses. Report from the Project on Children's Learning and Development. Technical Report No. 340. 137p. MF and HC available from EDRS.

MATHEMATICS EDUCATION RESEARCH STUDIES REPORTED IN JOURNALS AS INDEXED BY CURRENT INDEX TO JOURNALS IN EDUCATION

April - June 1976

EJ 128 414 Cox, L. S. Systematic Errors in the Four Vertical Algorithms in Normal and Handicapped Populations. Journal for Research in Mathematics Education, v6 n4, pp202-220, November 1975.

EJ 128 415 Sowder, Larry. Further Work on Verbalization of Discoveries. Journal for Research in Mathematics Education, v6 n4, pp222-227, November 1975.

EJ 128 416 Burton, John K.; And Others. Transfer Effects as a Function of Variety of Method and Ability Level in Elementary Mathematics. Journal for Research in Mathematics Education, v6 n4, pp228-231, November 1975.

EJ 128 417 Taloumis, Thalia. The Relationship of Area Conservation to Area Measurement as Affected by Sequence of Presentation of Piagetian Area Tasks to Boys and Girls in Grades One Through Three. Journal for Research in Mathematics Education, v6 n4, pp232-242, November 1975.

EJ 128 418 House, Peggy A. Learning Environments, Academic Self-Concepts, and Achievement in Mathematics. Journal for Research in Mathematics Education, v6 n4, pp244-252, November 1975.

EJ 128 419 Suydam, Marilyn N.; Weaver, J. F. Research on Mathematics Education (K-12) Reported in 1974. Journal for Research in Mathematics Education, v6 n4, pp253-282, November 1975.

EJ 130 043 Reys, Robert E.; Grouws, Douglas A. Division Involving Zero: Some Revealing Thoughts from Interviewing Children. School Science and Mathematics, v75 n7, pp593-605, November 1975.

EJ 131 416 Schafer, Pauline; And Others. Calculators in Some Fifth-Grade Classrooms: A Preliminary Look. Elementary School Journal, v76 n1, pp27-31, October 1975.

EJ 131 473 Scandura, Joseph M. How Does Mathematics Learning Take Place? Educational Studies in Mathematics, v6 n3, pp375-385, November 1975.

EJ 131 572 Yeshurun, Shraga. A Psychomathematically-Based Teaching Method. International Journal of Mathematical Education in Science and Technology, v6 n4, pp445-459, November 1975.

EJ 131 574 Allen, Layman E.; Ross, Joan. Instructional Gaming as a Means to Achieve Skill in Selecting Ideas Relevant for Solving a Problem. International Journal of Mathematical Education in Science and Technology, v6 n4, pp475-488, November 1975.